Box Seq. A



HOWREY & SIMON
Box No. 34
1299 Pennsylvania Avenue, N.W.
Washington, D.C. 20004-2402
(650) 463-8100

Attorney Docket No. 5371.31.US02

**Box Patent Application**ASSISTANT COMMISSIONER FOR PATENTS
Washington, D.C. 20231

Sir:

Transmitted herewith for filing is the patent application of Patricia D. Murphy, Marga B. White, Mark B. Rabin; Sheri J. Olson; Matthew Yoshikawa; Geoffrey M. Jackson; Tara Eskandari; Brenda Schryer; and Michael Park for NOVEL CODING SEQUENCE HAPLOTYPES OF THE HUMAN BRCA2 GENE.

Also, enclosed are:

- 1. Cover Sheet of Application;
- 2. 13 sheets of drawings;
- 3. Executed Declaration (2 sets);
- 4. Executed Small Entity Statement;
- 5. Assignment Cover Sheet;
- 6. Executed Assignment (2 sets);
- 7. Sequence Listing on Disk; and
- 5. Two (2) return postcards.

The filing fee has been calculated as shown below:

•	(Col. 1)	(	(Col. 2)
FOR	NO. FILED	NC	). EXTRA
BASIC FEE			
TOTAL CLAIMS	60-20 =	*	40
INDEP. CLAIMS	29 -3 =	*	26
■ MULTIPLE DEP	ENDENT CLAIN	1 PR	ESENTED

<sup>\*</sup>If the difference in Col. 1 is less than zero, enter "0" in Col. 2

SMALL	ENTITY	
RATE	FEE	
	\$ 395.00	
40 X 11 =	440.00	
26 x 41 =	1,066.00	
+ 135 =	135.00	
TOTAL	\$2,036.00	

	011121	THAN A LENTITY
OR	RATE	FEE
OR		\$ 790.00
OR	X 22 =	
OR	X 82=	
OR	- 270 =	
OR	TOTAL	

A check in the amount of \$2,076.00 (\$2,036.00 filing fee + \$40.00 recording of Assignment) is enclosed.

The U.S. Patent and Trademark Office is hereby authorized to charge any fee deficiency, or credit any overpayment, to our Deposit Account No. <u>08-3038</u>. A duplicate copy of this sheet is attached.

Date May 22, 1998

Albert P. Halluin (Reg. No.25,227)

PTO/SB/05 (2/98)
Please type a plus sign (+) inside this box + + Approved for use through 09/30/00 OMB 0651-0032
Patent and Trademark Office: U.S. DEPARTMENT OF COMMERCE
Under the provided Head of t

0 6
2.2
(O)
00

~			UTILITY		Attorney	Docke	t No.		5371.31.1	US02	
	2	PA	ATENT APPLICATIO	ON	First Nam	ned Inv	ento	r or Ap	plication	Identifier	Patricia D. Murphy
Ì	198	for new non	TRANSMITTAL provisional applications under	or 37 CFR 1 53(h))	Title	Nove Gene		ding S	Sequence	Haplotypes of the	he Human BRCA2
		T o			Express :			No.	EM5552	262526US	
			PPLICATION ELEMENT 600 concerning utility patent	· ·	· · · ·		Al	DDRE	ESS TO:	Assistant Com Box Patent Ap Washington, I	
	1.		smittal Form (Form PTC original and a duplicate for fe				6.		Microfie	che Computer Pr	rogram (Appendix)
	2.	Specificat			204 ]		7.			d/or Amino Acid	d Sequence Submission
		- Descript	ive title of the Invention eferences to Related App	lications				a.		omputer Readab	le Copy
		- Statemer	nt Regarding Fed sponsor	red R&D				b.	Pa	aper Copy (ident	ical to computer copy)
		- Backgro	und of the Invention mmary of the Invention					c.		atement verifyin	g identity of above
			scription of the Drawing	s (if filed)		Γ	ACC	OME		G APPLICATI	ON PARTS
Ī			_	s (II IIIcu)							
		- Claims	Description			ŀ	8. 9.		37 CFR :	3.73(b) Statemen	
1	İ_		of the Disclosure	[		į				ere is an assignee)	Attorney
- - -	3.	. Drawing(s	s) (35 USC 113)	[Total Sheets [		ı	10.	Ш	English	Translation Doc	rument (if applicable)
	4.	Oath or Declar	ation	[Total Pages	6 ]		11.			tion Disclosure nt (IDS)/P1O-1449	Copies of IDS Citations
æ		a. New	ly executed (original or c	сору)			12.		Prelimir	nary Amendmen	t
			py from a prior application for continuation/divisional with [Note Box 5 bd]	th Box 17 completed)			13.	$\overline{\boxtimes}$		Receipt Postcard e specifically itemiz	(MPEP 503) (Two) <i>sed)</i>
And the goal from the first			DELETION OF INVEN' Signed statement attached del in the prior application, see 3' 1.33(b)	TOR(S) leting inventor(s) named			14.	$\boxtimes$	*Small Stateme		Statement filed in prior application, Status still proper and desired
217							15.			d Copy of Prioris	ty Document(s)
	5.		tion By Reference (useal				16.		Other		,
			disclosure of the prior a			L				· · · · · · · · · · · · · · · · · · ·	
			h or declaration is supplicated of the disclosure of the disclosure of the			Г					ITLED TO PAY SMALL ENTITY
			eby incorporated by refer		cation						ED (37 C F R § 1 27), EXCEPT IF UPON (37 C F R § 1 28)
	17		UING APPLICATION,	•	and sum	nly the	reat	iisite i	informati	on.	
	` `	Continuatio		Continuation-in-part (			-		ication No		
			Prior Applica	ation Information: Exa	ımıner:		•	Grou	ıp/Art Un	iit:	
	<del> </del>			18. CORRESP		CE AI	DDR		· p/ 1 1 1 0 1		
	┢			100 00101001	01,122						
		Customer Numbe	r or Bar Code Label (Insc	ert Customer No or Attack	bar code l	abel hei	re)		or 🛛 (	Correspondence	address below
			Albert P. Hallum				· · ·		_		
	NA	ME	HOWREY & SIMON							••	
			Box No. 34	•							
	AD	DRESS	1299 Pennsylvania Ave	nue, N.W.							
	CII		Washington	STATE	DC	_			Z	IP CODE	20004-2402
	CO	UNTRY	US	TELEPHONE	202-78	33-080	00			4X	202-383-7195
	Nam	e (Print/Type)	Albert P. Hall		Registra	tion No	(Attor)	ney/Age	nt)	5 227	
		ature	Albert P. hall		1					25,227	

### APPLICATION IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

### **FOR**

### NOVEL CODING SEQUENCE HAPLOTYPES OF THE HUMAN BRCA2 GENE

### **INVENTORS**:

Patricia D. Murphy

U.S. citizen

Slingerlands, NY

Marga B. White

U.S. citizen

Frederick, MD

Mark B. Rabin

U.S. citizen

Rockville, MD

Sheri J. Olson

U.S. citizen

Falls Church, VA

Matthew Yoshikawa

U.S. citizen

Germantown, MD

Geoffrey M. Jackson

U.S. citizen

Beltsville, MD

Tara Eskandari

U.S. citizen

Rockville, MD

Brenda Schryer

U.S. citizen

Bel Air, MD

Michael Park

U.S. citizen

Rocksville, MD

HOWREY & SIMON 1299 Pennsylvania Ave., N.W. Box 34 Washington, D.C. 20004 (650) 463-8100

Attorney Docket No. 5371.31.US02

### NOVEL CODING SEQUENCE HAPLOTYPES OF THE HUMAN BRCA2 GENE

This is an U.S. utility patent application based on U.S. Provisional Application Serial Nos. 60/055,784 filed on August 15, 1997, 60/064,926 filed on November 7, 1997, and 60/065,367 filed on November 12, 1997.

### FIELD OF THE INVENTION

This invention relates to a gene which has been associated with breast cancer where the gene is found to be mutated. More specifically, this invention relates to five unique coding sequences of BRCA2 gene BRCA2<sup>(omi1)</sup>, BRCA2<sup>(omi2)</sup>, BRCA2<sup>(omi3)</sup>, BRCA2<sup>(omi4)</sup>, and BRCA2<sup>(omi5)</sup> identified in human subjects which define five novel haplotypes.

### BACKGROUND OF THE INVENTION

It has been estimated that about 5-10% of breast cancer is inherited (Rowell, S., et al., American Journal of Human Genetics 55:861-865 (1994)). The first gene associated with both breast and ovarian cancer was cloned in 1994 from chromosome 17 by Miki, Y., et al., Science 266:66-71 (1994). A second high-risk breast cancer conferring gene was located on chromosome 13 in 1994 (Wooster, R., et al., Science 265:2088-2090) and subsequently cloned in 1995 (Wooster, R., et al., Nature 378:789-792). Mutations in this "tumor suppressor" gene are thought to account for roughly 35% of inherited breast cancer and 80-90% of families with male breast cancer.

Locating one or more mutations in the BRCA2 region of chromosome 13 provides a promising approach to reducing the high incidence and mortality associated with breast cancer through the early detection of women and men at high risk. These individuals, once identified, can be targeted for more aggressive prevention programs. Screening is carried out by a variety of methods which include karyotyping, probe binding and DNA sequencing.

In DNA sequencing technology, genomic DNA is extracted from whole blood and the coding regions of the BRCA2 gene are amplified. Each of the coding regions may be sequenced completely and the results are compared to the normal DNA sequence of the gene. Alternatively, the coding sequence of the sample gene may be compared to a panel of known mutations or other screening procedure

before completely sequencing the gene and comparing it to a normal sequence of the gene.

The BRCA2 gene is divided into 27 separate exons. Exon 1 is noncoding, in that it is not part of the final functional BRCA2 protein product. The BRCA2 coding region spans roughly 10433 base pairs (bp) over 70 kb. Each exon consists of 100-600 bp, except for exons 10, 11 and 27. The full length mRNA is 11-12 kb. To sequence the coding region of the BRCA2 gene, each exon is amplified separately and the resulting PCR products are sequenced in the forward and reverse directions. Because exons 10, 11, and 27 are so large, we have divided them into three, twenty-one, and two overlapping PCR fragments (respectively) of approximately 250-625 bp each (segments "A" through "C" of exon 10, "A" through "U" of exon 11, and "A" through "B" of exon 27).

Many mutations and normal polymorphisms have already been reported in the BRCA2 gene. A world wide web site has been built to facilitate the detection and characterization of alterations in breast cancer susceptibility genes. Such mutations in BRCA2 can be accessed through the Breast Cancer Information Core (BIC) at http://www.nhgri.nih.gov/Intramural\_research/Lab\_transfer/Bic. This data site became publicly available on November 1, 1995. Friend, S. et al. Nature Genetics 11:238, (1995). The information on BRCA2 was added in February, 1996.

The genetics of Breast Cancer Syndrome is autosomal dominant with reduced penetrance. In simple terms, this means that the syndrome runs through families: (1) both sexes can be carriers (mostly women get the disease but men can both pass it on and occasionally get breast cancer); (2) most generations will likely have breast cancer; (3) occasionally women carriers either die young before they have the time to manifest disease (and yet have offspring who get it) or they never develop breast or ovarian cancer and die of old age (the latter people are said to have "reduced penetrance" because they never develop cancer). Pedigree analysis and genetic counseling is absolutely essential to the proper workup of a family prior to any lab work.

Until now, the only sources of genomic sequence information for BRCA2 were GenBank (Accession Number U43746), or through the Breast Information Core (BIC) database on the Internet which requires membership in the BIC consortium. However, based upon the disclosure of this patent application, in neither GenBank

nor BIC were the sequences identified and listed entirely accurate. There is a need in the art to correct these mistakes which otherwise may lead to misinterpretation of the sequence data from the patient as abnormal when it was not, or vice versa.

In addition, there is a need in the art to have available a functional allele profile which represents the most likely BRCA2 sequences to be found in the majority of the normal population. This functional allele profile is based upon frequent polymorphisms and the correct backbone sequence. The knowledge of several common normal haplotypes will make it possible for true mutations to be easily identified or differentiated from polymorphisms. Identification of mutations of the BRCA2 gene and protein would allow more widespread diagnostic screening for hereditary breast cancer than is currently possible.

The use of these common normal haplotypes, in addition to the previously published BRCA2 sequence, will reduce the likelihood of misinterpreting a "sequence variation" found in the normal population with a pathologic "mutation" (i.e. causes disease in the individual or puts the individual at a high risk of developing the disease). With large interest in breast cancer predisposition testing, misinterpretation is particularly worrisome. People who already have breast cancer are asking the clinical question: "is my disease caused by a heritable genetic mutation?" The relatives of the those with breast cancer are asking the question: "Am I also a carrier of the mutation my relative has? Thus, is my risk increased, and should I undergo a more aggressive surveillance program?"

### **SUMMARY OF THE INVENTION**

The present invention is based on the discovery of the correct genomic BRCA2 sequence and five novel sequence haplotypes found in normal human subjects of the BRCA2 gene.

It is an object of this invention to provide the correct intronic/exonic sequence of the BRCA2 gene.

It is another object of this invention to provide five unique haplotype sequences of the BRCA2 gene in normal individuals which do not correspond to increased cancer susceptibility.

It is another object of this invention to sequence a BRCA2 gene or a portion thereof and compare it to the five haplotype sequences to determine whether a

sequence variation noted represents a polymorphism or a potentially harmful mutation.

It is another object of this invention to provide a list of the pairs which occur at each of ten polymorphic points in the BRCA2 gene.

It is another object of this invention to provide the rates of occurrence for the polymorphisms at codons 289, 372, 455, 743, 894, 991, 1132, 1269, 2414, and 2951 in the BRCA2 gene.

It is another object of this invention to provide a method wherein all exons of BRCA2 gene or parts thereof, are amplified with one or more oligonucleotide primers.

It is another object of this invention to provide a method of identifying a individual who carries no mutation(s) of the BRCA2 gene and is therefore at no increased risk or susceptibility to breast or ovarian cancer based on a finding that the individual does not carry an abnormal BRCA2 genes.

It is another object of this invention to provide a method of identifying a mutation in BRCA2 gene leading to predisposition or higher susceptibility to breast or ovarian cancer.

It is another object of this invention to provide five novel BRCA2 protein sequences derived from five BRCA2 haplotype sequences.

It is another object of the invention to encompass prokaryotic or eukaryotic host cells comprising an expression vector having a DNA sequence that encodes for all or a fragment of the five novel BRCA2 protein sequences, a BRCA2 polypeptide thereof, or a functional equivalent thereof.

It is another object of the invention to encompass an anti-BRCA2 protein antibody using all of fragments of the five novel BRCA2 protein sequences, a BRCA2 polypeptide thereof or a functional equivalent thereof as an immunogen.

There is a need in the art for cDNA sequences of the BRCA2 gene and for the protein sequences of BRCA2 gene from normal individuals who are not at risk for increased susceptibility for cancer. In order to determine whether a sample from a patient suspected of containing a BRCA2 mutation actually has the mutation, the patient's BRCA2 DNA and/or amino acid sequence need to be compared to all known normal BRCA2 sequences. Failure to compare the sequence obtained to all

naturally occurring normal sequences may result in reporting a sample as containing a potentially harmful mutation when it is a polymorphism without clinical significance.

A person skilled in the art of genetic susceptibility testing will find the present invention useful for:

- identifying individuals having a normal BRCA2 gene with no coding sequence mutations, who therefore cannot be said to have an increased genetic susceptibility to breast or ovarian cancer from their BRCA2 genes;
- avoiding misinterpretation of normal polymorphisms found in the BRCA2 gene;
- determining the presence of a previously unknown mutation in the BRCA2 gene;
- d) identifying a mutation in exon 11 of BRCA2 which indicates a predisposition or higher susceptibility to ovarian cancer than breast cancer (i.e., resides in the putative "ovarian cancer cluster" region);
- e) probing a human sample of the BRCA2 gene by allele to determine the presence or absence of either polymorphic alleles or mutations;
- f) performing gene therapy with the correct BRCA2 gene sequence.
- g) performing protein replacement therapy with the correct BRCA 2 protein sequence or a functional equivalent thereof.

### **BRIEF DESCRIPTION OF THE FIGURES**

FIGURE 1 shows the GenBank genomic sequence of BRCA2 (Accession Number U43746). The lower case letters denote intronic sequences and the upper case letters denote exonic sequences. Incorrect exonic sequences at exons 5 and 16 are shown with boldface type.

FIGURE 2 shows the corrected genomic sequence of BRCA2. The lower case letters denote intronic sequences and the upper case letters denote exonic sequences. Corrected intronic and exonic sequences at exons 5, 11 and 15 are shown with boldface type.

FIGURE 3 shows the alternative alleles at polymorphic sites along a chromosome which can be represented as a unit or "haplotype" within a gene such as BRCA2.

The haplotype that is in GenBank (GB) is shown with light shading. Five additional haplotypes are shown in FIGURE 3 (encompassing the alternative alleles found at nucleotide sites 1093, 1342, 1593, 2457, 2908, 3199, 3624, 4035, 7470 and 9079). BRCA2 (omi-1), BRCA2 (omi-2), BRCA2 (omi-3), BRCA2 (omi-4), and BRCA2 (omi-5) are represented with mixed dark and light shading (numbers 2, 4, 6, 8 and 10 from left to right). In total, 5 of 10 haplotypes along the BRCA2 gene are unique.

### **DETAILED DESCRIPTION OF THE INVENTION**

### **DEFINITIONS**

The following definitions are provided for the purpose of understanding this invention.

"Breast and Ovarian cancer" is understood by those skilled in the art to include breast, ovarian and pancreatic cancer in women and also breast, prostate and pancreatic cancer in men. BRCA2 is associated with genetic susceptibility to breast, ovarian and pancreatic cancer. Therefore, claims in this document which recite breast and/or ovarian cancer refer to breast, ovarian, prostate, and pancreatic cancers in men and women.

"Coding sequence" refers to those portions of a gene which, taken together, code for a peptide (protein), or which nucleic acid itself has function.

"Protein" or "peptide" refers to a sequence of amino acids which has function.

"BRCA2<sup>(omi)</sup>" refers to the genomic BRCA2 sequence disclosed in Genbank (Accession Number U43746) wherein,

- (1) a 10 bp stretch (5'-TTTATTTTAG-3') is intronic at 3' end of intron 4, rather than at the 5' end of exon 5; and
- (2) a 16 bp stretch (5'-GTGTTCTCATAAACAG-3') is exonic at the 3' end of exon 15, rather than at the 5' end of exon.

"BRCA2<sup>(omi 1-5)</sup>" refers to five unique DNA sequences of the BRCA2 gene and their introns (particularly the slice sites adjacent to the exons). These sequences were found by end to end sequencing of the BRCA2 gene from 5 individuals randomly drawn from the population and who were documented to have no family history of breast or ovarian cancer. The sequenced exons were found not to contain any truncating mutations. In all cases the change of a nucleic acid at a

polymorphic site lead to a codon change and a change of amino acid from the previously published standard in GenBank (see TABLE III). In some cases the frequency of occurrence of a nucleic acid change was found to differ from the published frequency or was newly determined. These sequence variations are believed to be alleles whose haplotypes do not indicate an increased risk for cancer.

"Normal DNA sequence" also called "normal gene sequence" refers to a nucleic acid sequence, the nucleic acid of which are known to occur at their respective positions with high frequency in a population of individuals who carry the gene which codes for a normally functioning protein, or which itself has normal function.

"Normal Protein Sequence" refers to the protein sequence, the amino acids of which are known to occur with high frequency in a population of individuals who carry the gene which codes for a normally functioning protein.

"Normal Sequence" refers to the nucleic acid or protein sequence, the nucleic or amino acids of which are known to occur with high frequency in a population of individuals who carry the gene which codes for a normally functioning protein, or which nucleic acid itself has a normal function.

"Haplotype" refers to a series of specific alleles within a gene along a chromosome.

"Functional allele profile" refers a list of those alleles in the normal population which have the funll function.

"Mutation" refers to a base change or a gain or loss of base pair(s) in a DNA sequence, which results in a DNA sequence coding for a non-functional protein or a protein with substantially reduced or altered function.

"Polymorphism" refers to a base change in a DNA sequence which is not associated with known pathology.

"Primer" refers to a sequence comprising about 15 or more nucleotides having a sequence complementary to the BRCA2 gene. Other primers which can be used for primer hybridization will be known or readily ascertainable to those skilled in the art.

"Substantially complementary to" refers to primer sequences which hybridize to the sequences provided under stringent conditions and/or sequences having

sufficient homology with BRCA2 sequences, such that the allele specific oligonucleotide primers hybridize to the BRCA2 sequences to which they are complimentary.

"Isolated nucleic acids" refers to nucleic acids substantially free of other nucleic acids, proteins, lipids, carbohydrates or other materials with which they may be associated. Such association is typically either in cellular material or in a synthesis medium.

"Biological sample" or "body sample" refers to a sample containing DNA oatained from a biological source. The sample may be from a living, dead or even archeological source from a variety of tissues and cells. Examples include body fluid (e.g. blood (leukocytes), urine (epithelial cells), saliva, breast milk, menstrual flow, cervical and vaginal secretions, etc.), skin, hair roots/follicle, mucus membrane (e.g. buccal or tongue cell scrapings), cervicovaginal cells (from PAP smear, etc.), lymphatic tissue, internal tissue (normal or tumor).

"Vector" refers to any polynucleotide which is capable of self replication or inducing integration into a self-replicating polynucleotide. Examples include polynucleotides containing an origin or replication or an integration site. Vectors may be intergrated into the host cell's chromosome or form an autonomously replicating unit.

"A tumor growth inhibitor" refers to a molecule such as, all or a fragment of BRCA2 protein, a BRCA2 polypeptide, or a functional equivalent thereof that is effective for preventing the formation of, reducing, or eliminating a transformed or malignant phenotype of breast or ovarian cancer cells.

"A BRCA2 polypeptide" refers to a BRCA2 polypeptide either directly derived from the BRCA2 protein, or homologous to the BRCA2 protein, or a fusion protein consisting of all or fragments of the BRCA2 protein and polypeptides.

"A functional equivalent" refers to a molecule including an unnatural BRCA2 polypeptide, a drug or a natural product which retains substantial biological activity as the native BRCA2 protein. The activity and function of BRCA2 protein may include transactivation, granin, DNA repair, among others.

"A target polynucleotide" refers to the nucleic acid sequence of interest, for example, the BRCA2 encoding polynucleotide. Other primers which can be used for primer hybridization will be known or readily ascertainable to those of skill in the art.

The invention in several of its embodiments includes: an isolated DNA sequence of the BRCA2 coding sequence as set forth in SEQ ID NO:4, 6, 8, 10, and 12, a protein sequence of the BRCA2 protein as set forth in SEQ ID NO:5, 7, 9, 11, 13, a method of identifying individuals having a normal BRCA2 gene with no increased risk for breast and ovarian cancer, a method of detecting an increased genetic susceptibility to breast and ovarian cancer in an individual resulting from the presence of a mutation in the BRCA2 coding sequence, a method of performing gene therapy to prevent or treat a tumor, a method of protein replacement therapy to prevent or treat a tumor, a diagnostic reagent comprising all or fragments of the disclosed BRCA2 cDNA and protein sequences.

### SEQUENCING

Any nucleic acid specimen, in purified or non-purified form, can be utilized as the starting nucleic acid, providing it contains, or is suspected of containing, the specific nucleic acid sequence containing a polymorphic or a mutant allele. Thus, the process may amplify, for example, DNA or RNA, including mRNA and cDNA, wherein DNA or RNA may be single stranded or double stranded. In the event that RNA is to be used as a template, enzymes and/or conditions optimal for reverse transcribing the template to DNA would be utilized. In addition, a DNA-RNA hybrid which contains one strand of each may be utilized. A mixture of nucleic acids may also be employed, or the nucleic acids produced in a previous method such as an amplification reaction using the same or different primers may be so utilized. The specific nucleic acid sequence to be amplified, i.e., the polymorphic and/or the mutant allele, may be a fraction of a larger molecule or can be present initially as a discrete molecule, so that the specific sequence constitutes the entire nucleic acid. A variety of amplification techniques may be used such as ligating the DNA sample or fragments thereof to a vector capable of replication or incorporation into a replicating system thereby increasing the number of copies of DNA suspected of containing at least a portion of the BRCA2 gene. Amplification techniques include so called "shot gun cloning". It is not necessary that the sequence to be amplified be present initially in a pure form; it may be a minor fraction of a complex mixture, such as contained in whole human DNA.

It should be noted that one need not sequence the entire coding region or even an entire DNA fragment in order to determine whether or not a mutation is present. For example, when a mutation is known in one family member, it is sufficient to determine the sequence at only the mutation site by sequencing or by other mutation detection systems such as ASO when testing other family members.

DNA utilized herein may be extracted from a body sample, such as blood, tissue material and other biological sample by a variety of techniques such as that described by Maniatis, et al. in Molecular Cloning: A Laboratory Manual, Cold Spring Harbor, NY, p 280-281, 1982). If the extracted sample is impure, it may be treated before amplification with an amount of a reagent effective to open the cells, and to expose and/or separate the strand(s) of the nucleic acid(s). This lysing and nucleic acid denaturing step to expose and separate the strands will allow amplification to occur much more readily.

For amplification by cloning, the isolated DNA may be cleaved into fragments by a restriction endonuclease or by shearing by passing the DNA containing mixture through a 25 gauge needle from a syringe to prepare 1-1.5 kb fragments. The fragments are then ligated to a cleaved vector (virus, plasmid, transposon, cosmid etc.) and then the recombinant vector so formed is then replicated in a manner typical for that vector.

For a PCR amplification, the deoxyribonucleotide triphosphates dATP, dCTP, dGTP, and dTTP are added to the synthesis mixture, either separately or together with the primers, in adequate amounts and the resulting solution is heated to about 90°-100°C from about 1 to 10 minutes, preferably from 1 to 4 minutes. After this heating period, the solution is allowed to cool, which is preferable for the primer hybridization. To the cooled mixture is added an appropriate agent for effecting the primer extension reaction (called herein "agent for polymerization"), and the reaction is allowed to occur under conditions known in the art. The agent for polymerization may also be added together with the other reagents if it is heat stable. This synthesis (or amplification) reaction may occur at room temperature up to a temperature above which the agent for polymerization no longer functions. Thus, for example, if DNA polymerase is used as the agent, the temperature is generally no greater than about 40°C. Most conveniently the reaction occurs at

room temperature. When using thermostable DNA polymerase such as Taq, higher temperature may be used.

The allele specific oligonucleotide primers are useful in determining whether a subject is at risk of having breast or ovarian cancer, and also useful for characterizing a tumor. Primers direct amplification of a target polynucleotide prior to sequencing. These unique BRCA2 oligonucleotide primers for exons 2-27 shown in TABLE II were designed and produced specifically to optimize amplification of portions of BRCA2 which are to be sequenced.

The primers used to carry out this invention embrace oligonucleotides of sufficient length and appropriate sequence to provide initiation of polymerization. Environmental conditions conducive to synthesis include the presence of nucleoside triphosphates and an agent for polymerization, such as DNA polymerase, and a suitable temperature and pH. The primer is preferably single stranded for maximum efficiency in amplification, but may be double stranded. If double stranded, the primer is first treated to separate its strands before being used to prepare extension products. The primer must be sufficiently long to prime the synthesis of extension products in the presence of the inducing agent for polymerization. The exact length of primer will depend on many factors, including temperature, buffer, and nucleotide composition. The oligonucleotide primer typically contains 18-28 bp plus in some cases an M13 "tail" for convenience.

Primers used to carry out this invention are designed to be substantially complementary to each strand of the genomic locus to be amplified. This means that the primers must be sufficiently complementary to hybridize with their respective strands under conditions which allow the agent for polymerization to perform. In other words, the primers should have sufficient complementarity with the 5' and 3' sequences flanking the mutation to hybridize therewith and permit amplification of the genomic locus.

Oligonucleotide primers of the invention are employed in the amplification process which is an enzymatic chain reaction that produces exponential quantities of polymorphic locus relative to the number of reaction steps involved. Typically, one primer is complementary to the negative (-) strand of the polymorphic locus and the other is complementary to the positive (+) strand. Annealing the primers to denatured nucleic acid followed by extension with an enzyme, such as the large

fragment of DNA polymerase I (Klenow) and nucleotides, results in newly synthesized + and - strands containing the target polymorphic locus sequence. Because these newly synthesized sequences are also templates, repeated cycles of denaturing, primer annealing, and extension results in exponential production of the region (*i.e.*, the target polymorphic locus sequence) defined by the primers. The product of the chain reaction is a discreet nucleic acid duplex with termini corresponding to the ends of the specific primers employed.

The oligonucleotide primers of the invention may be prepared using any suitable method, such as conventional phosphotriester and phosphodiester methods or automated embodiments thereof. In one such automated embodiment, diethylphosphoramidites are used as starting materials and may be synthesized as described by Beaucage, et al., Tetrahedron Letters, 22:1859-1862, 1981. One method for synthesizing oligonucleotides on a modified solid support is described in U.S. Patent No. 4,458,066.

The agent for polymerization may be any compound or system which will function to accomplish the synthesis of primer extension products, including enzymes. Suitable enzymes for this purpose include, for example, *E. coli* DNA polymerase I, Klenow fragment of *E. coli* DNA polymerase, polymerase muteins, reverse transcriptase, other enzymes, including heat-stable enzymes (*i.e.*, those enzymes which perform primer extension after being subjected to temperatures sufficiently elevated to cause denaturation), such as *Taq* polymerase. Suitable enzymes will facilitate combination of the nucleotides in the proper manner to form the primer extension products which are complementary to each polymorphic locus nucleic acid strand. Generally, the synthesis will be initiated at the 3' end of each primer and proceed in the 5' direction along the template strand, until synthesis terminates, producing molecules of different lengths.

The newly synthesized strand and its complementary nucleic acid strand will form a double-stranded molecule under hybridizing conditions described above and this hybrid is used in subsequent steps of the process. In the next step, the newly synthesized double-stranded molecule is subjected to denaturing conditions using any of the procedures described above to provide single-stranded molecules.

The steps of denaturing, annealing, and extension product synthesis can be repeated as often as needed to amplify the target polymorphic locus nucleic acid

sequence to the extent necessary for detection. The amount of the specific nucleic acid sequence produced will accumulate in an exponential fashion. Amplification is described in *PCR*. *A Practical Approach*, ILR Press, Eds. M. J. McPherson, P. Quirke, and G. R. Taylor, 1992.

The amplification products may be detected by Southern blots analysis, without using radioactive probes. In such a process, for example, a small sample of DNA containing a very low level of the nucleic acid sequence of the polymorphic locus is amplified, and analyzed via a Southern blotting technique or similarly, using dot blot analysis. The use of non-radioactive probes or labels is facilitated by the high level of the amplified signal. Alternatively, probes used to detect the amplified products can be directly or indirectly detectably labeled, for example, with a radioisotope, a fluorescent compound, a bioluminescent compound, a chemiluminescent compound, a metal chelator or an enzyme. Those of ordinary skill in the art will know of other suitable labels for binding to the probe, or will be able to ascertain such, using routine experimentation.

Sequences amplified by the methods of the invention can be further evaluated, detected, cloned, sequenced, and the like, either in solution or after binding to a solid support, by any method usually applied to the detection of a specific DNA sequence such as PCR, oligomer restriction (Saiki, et.al., Bio/Technology, 3:1008-1012, 1985), allele-specific oligonucleotide (ASO) probe analysis (Conner, et al., Proc. Natl. Acad. Sci. U.S.A., 80:278, 1983), oligonucleotide ligation assays (OLAs) (Landgren, et al., Science, 241:1007, 1988), and the like. Molecular techniques for DNA analysis have been reviewed (Landgren, et al., Science, 242:229-237, 1988).

Preferably, the method of amplifying is by PCR, as described herein and as is commonly used by those of ordinary skill in the art. Alternative methods of amplification have been described and can also be employed as long as the BRCA2 locus amplified by PCR using primers of the invention is similarly amplified by the alternative means. Such alternative amplification systems include but are not limited to self-sustained sequence replication, which begins with a short sequence of RNA of interest and a T7 promoter. Reverse transcriptase copies the RNA into cDNA and degrades the RNA, followed by reverse transcriptase polymerizing a second strand of DNA. Another nucleic acid amplification technique

is nucleic acid sequence-based amplification (NASBA) which uses reverse transcription and T7 RNA polymerase and incorporates two primers to target its cycling scheme. NASBA can begin with either DNA or RNA and finish with either, and amplifies to 10<sup>8</sup> copies within 60 to 90 minutes. Alternatively, nucleic acid can be amplified by ligation activated transcription (LAT). LAT works from a singlestranded template with a single primer that is partially single-stranded and partially double-stranded. Amplification is initiated by ligating a cDNA to the promoter oligonucleotide and within a few hours, and amplification is  $10^8\,$  to  $10^9\,$  fold. Another amplification system useful in the method of the invention is the  $\ensuremath{\text{Q}\beta}$ Replicase System. The Qβ replicase system can be utilized by attaching an RNA sequence called MDV-1 to RNA complementary to a DNA sequence of interest. Upon mixing with a sample, the hybrid RNA finds its complement among the specimen's mRNAs and binds, activating the replicase to copy the tag-along sequence of interest. Another nucleic acid amplification technique, ligase chain reaction (LCR), works by using two differently labeled halves of a sequence of interest which are covalently bonded by ligase in the presence of the contiguous sequence in a sample, forming a new target. The repair chain reaction (RCR) nucleic acid amplification technique uses two complementary and target-specific oligonucleotide probe pairs, thermostable polymerase and ligase, and DNA nucleotides to geometrically amplify targeted sequences. A 2-base gap separates the oligonucleotide probe pairs, and the RCR fills and joins the gap, mimicking normal DNA repair. Nucleic acid amplification by strand displacement activation (SDA) utilizes a short primer containing a recognition site for hincll with short overhang on the 5' end which binds to target DNA. A DNA polymerase fills in the part of the primer opposite the overhang with sulfur-containing adenine analogs. HincII is added but only cuts the unmodified DNA strand. A DNA polymerase that lacks 5' exonuclease activity enters at the site of the nick and begins to polymerize, displacing the initial primer strand downstream and building a new one which serves as more primer. SDA produces greater than 10<sup>7</sup>-fold amplification in 2 hours at 37°C. Unlike PCR and LCR, SDA does not require instrumented Temperature cycling.

Another method is a process for amplifying nucleic acid sequences from a DNA or RNA template which may be purified or may exist in a mixture of nucleic acids. The resulting nucleic acid sequences may be exact copies of the template, or may be modified. The process has advantages over PCR in that it increases the fidelity of copying a specific nucleic acid sequence, and it allows one to more efficiently detect a particular point mutation in a single assay. A target nucleic acid is amplified enzymatically while avoiding strand displacement. Three primers are used. A first primer is complementary to the first end of the target. A second primer is complementary to the second end of the target. A third primer which is similar to the first end of the target and which is substantially complementary to at least a portion of the first primer such that when the third primer is hybridized to the first primer, the position of the third primer complementary to the base at the 5' end of the first primer contains a modification which substantially avoids strand displacement. This method is detailed in U.S. Patent 5,593,840 to Bhatnagar et al. 1997, incorporated herein by reference.

Finally, recent application of DNA chips or microarray technology where DNA or oligonucleotides are immobilized on small solid support may also be used to rapidly sequence sample BRCA2 gene and analyze its expression. Typically, high density arrays of DNA fragment are fabricated on glass or nylon substrates by *in situ* light-directed combinatorial synthesis or by conventional synthesis followed by immobilization (Fodor *et al.* U.S. patent No. 5,445,934). Sample DNA or RNA may be amplified by PCR, labeled with a fluorescent tag, and hybridized to the microarray. Examples of this technology are provided in U.S. Patents 5,510, 270, U.S. 5,547,839, incorporated herein by reference.

All exonic and adjacent intronic sequences of the BRCA2 gene were obtained by end to end sequencing of five normal subjects in the manner described above followed by analysis of the data obtained. The data obtained provided us with the opportunity to establish the correct intronic/exonic structure of the BRCA2 gene. In addition, we evaluated six previously published normal polymorphisms (1342, 2457, 3199, 3624, 4035, and 7470) for correctness and frequency in the population, and to identify four additional polymorphisms not previously characterized (1093, 1593, 2908, and 9079).

### **GENE THERAPY**

The polynucleotide(s) which result from either sense or antisense transcription of any exon or the entire coding sequence or fragments of BRCA2 gene may be used for gene therapy. A variety of methods are known for gene transfer, any of which might be available for use.

Direct injection of Recombinant DNA in vivo:

- 1. Direct injection of "naked" DNA directly with a syringe and needle into a specific tissue, infused through a vascular bed, or transferred through a catheter into endothelial cells.
- 2. Direct injection of DNA that is contained in artificially generated lipid vesicles or other encapsulating vehicles.
- 3. Direct injection of DNA conjugated to a target receptor structure, such as a diptheria toxin, an antibody or other suitable receptor.
- 4. Direct injection by particle bombardment. For example, the DNA may be coated onto gold particles and shot into the cells.

### **Human Artificial Chromosomes**

The gene delivery approach involves the use of human chromosomes that have been stripped down to contain only the essential components for replication and the genes desired for transfer.

### Receptor-Mediated Gene Transfer

DNA is linked to a targeting molecule that will bind to specific cell-surface receptors, inducing endocytosis and transfer of the DNA into mammalian cells. One such technique uses poly-L-lysine to link asialoglycoprotein to DNA. An adenovirus is also added to the complex to disrupt the lysosomes and thus allow the DNA to avoid degradation and move to the nucleus. Infusion of these particles intravenously has resulted in gene transfer into hepatocytes.

### RECOMBINANT VIRUS VECTORS

Several vectors may be used in gene therapy. Among them are the Moloney Murine Leukemia Virus (MoMLV) Vectors, the adenovirus vectors, the Adeno-Associated Virus (AAV) vectors, the herpes simplex virus (HSV) vectors, the poxvirus vectors, the retrovirus vectors, and human immunodeficiency virus (HIV) vectors.

### GENE REPLACEMENT AND REPAIR

The ideal genetic manipulation for treatment of a genetic disease would be the actual replacement of the defective gene with a normal copy of the gene. Homologous recombination is the term used for switching out a section of DNA and replacing it with a new piece. By this technique, the defective gene may be replaced with a normal gene which expresses a functioning BRCA2 tumor growth inhibitor protein.

A complete description of gene therapy can also be found in "Gene Therapy A Primer For Physicians" 2d Ed. by Kenneth W. Culver, M.D. Publ. Mary Ann Liebert Inc. (1996). Two Gene Therapy Protocols for BRCA1 gene have been approved by the Recombinant DNA Advisory Committee for Jeffrey T. Holt et al. They are listed as 9602-148, and 9603-149 and are available from the NIH. Protocols for BRCA2 gene therapy may be similarly employed. The isolated BRCA2 gene may be synthesized or constructed from amplification products and inserted into a vector such as the LXSN vector.

### A BRCA2 POLYPEPTIDE OR ITS FUNCTIONAL EQUIVALENT

The growth of breast and ovarian cancer may be arrested or prevented by directly increasing the BRCA2 protein level where inadequate functional BRCA2 activity is responsible for breast and ovarian cancer. The cDNA and amino acid sequences of five novel BRCA2 haplotypes are disclosed herein (SEQ ID No:4-13). All or a fragment of BRCA2 protein may be used in therapeutic or prophylactic treatment of breast and ovarian cancer. Such a fragment may have a similar biological function as the native BRCA2 protein or may have a desired biological function as specified below. BRCA2 polypeptides or their functional equivalents including homologous and modified polypeptide sequences are also within the scope of the present invention. Changes in the native sequence may be advantageous in producing or using the BRCA2 derived polypeptides or functional equivalents suitable for therapeutic or prophylactic treatment of breast and ovarian cancer. For example, these changes may be desirable for producing resistance against *in vivo* proteolytic cleavage, for facilitating transportation and delivery of

therapeutic reagents, for localizing and compartmentalizing tumor suppressing agents, or for expression, isolating and purifying the target species.

There are a variety of methods to produce an active BRCA2 polypeptide or a functional equivalent as a tumor growth inhibitor. For example, one or more amino acids may be substituted, deleted, or inserted using methods well known in the art (Maniatis et al., 1982). Considerations of polarity, charge, solubility, hydrophobicity, hydrophilicity and/or the amphiphathic nature of the amino acids play an important role in designing homologous polypeptide changes suitable for the intended treatment. In particular, conservative amino acid substitution using amino acids that are related in side-chain structure and charge may be employed to preserve the chemical and biological property. A homologous polyeptide typically contains at least 70% homology to the native sequence. Unnatural forms of the polypeptide may also be incorporated so long as the modification retains substantial biological activity. These unnatural polypeptides typically include structural mimics and chemical medications, which have similar three-dimensional structures as the active regions of the native BRCA2 protein. For example, these modifications may include terminal D-amino acids, cyclic peptides, unnatural amino acids side chains, pseudopeptide bonds, N-terminal acetylation, glycosylation, and biotinylation, etc. These unnatural forms of polypeptide may have a desired biological function, for example, they may be particularly robust in the presence of cellular or serum proteases and exopeptidase. An effective BRCA2 polypeptide or a functional equivalent may also be recognized by the reduction of the native BRCA2 protein. Regions of the BRCA2 protein may be systematically deleted to identify which regions are essential for tumor growth inhibitor activity. These smaller fragments of BRCA2 protein may then be subjected to structural and functional modification to derive therapeutically or prophylactically effective regiments. Finally, drugs, natural products or small molecules may be screened or synthesized to mimic the function of the BRCA2 protein. Typically, the active species retain the essential threedimensional shape and chemical reactivity, and therefore retain the desired aspects of the biological activity of the native BRCA2 protein. The activity and function of BRCA2 may include transactivation, granin, DNA repair among others. Functions of BRCA2 protein are also reviewed in Bertwistle and Ashworth, Curr. Opin. Genet. Dev. 8(1): 14-20 (1998) and Zhang et al., Cell 92:433-436 (1998). It will be

apparent to one skilled in the art that a BRCA2 polypeptide or a functional equivalent may be selected because such polypeptide or functional equivalent possesses similar biological activity as the native BRCA2 protein.

### EXPRESSION OF THE BRCA2 PROTEIN AND POLYPEPTIDE IN HOST CELLS

All or fragments of the BRCA2 protein and polypeptide may be produced by host cells that are capable of directing the replication and the expression of foreign genes. Suitable host cells include prokaryotes, yeast cells, or higher eukaryotic cells, which contain an expression vector comprising all or a fragment of the BRCA2 cDNA sequence (SEQ. ID No: 4, 6, 8, 10, or 12) operatively linked to one or more regulatory sequences to produce the intended BRCA2 protein or polypeptide. Prokaryotes may include gram negative or gram positive organisms, for example *E. coli* or *Bacillus* strains. Suitable eukaryotic host cells may include yeast, virus, and mamalian systems. For example, Sf9 insect cells and human cell lines, such as COS, MCF7, HeLa, 293T, HBL100, SW480, and HCT116 cells.

A broad variety of suitable expression vectors are available in the art. An expression vector typically contains an origin of replication, a promoter, a phenotypic selection gene (antibiotic resistance or autotrophic requirement), and a DNA sequence coding for all or fragments of the BRCA2 protein. The expression vectors may also include other operatively linked regulatory DNA sequences known in the art, for example, stability leader sequences, secretory leader sequences, restriction enzyme cleavage sequences, polyadenylation sequences, and termination sequences, among others. The essential and regulatory elements of the expression vector must be compatible with the intended host cell. Suitable expression vectors containing the desired coding and control regions may be constructed using standard recombinant DNA techniques known in the art, many of which are described in Sambrook, et al., Molecular Cloning: A Laboratory Manual, Second Edition, Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y. (1989). For example, suitable origins of replication may include Col E1, SV4O viral and M13 origins of replication. Suitable promoters may be constitutive or inducible, for example, tac promoter, lac Z promoter, SV40 promoter, MMTV promoter, and LXSN promoter. Examples of selectable markers include neomycin, ampicillin, and hygromycin resistance and the like. Many suitable prokaryotic, viral and mammalian expression vectors may be obtained commercially, for example, from Invitrogen Corp., San Diego, CA or from Clontech, Palo Alto, CA. It may be desirable that the BRCA2 protein or polypeptide is produced as a fusion protein to enhance the expression in selected host cells, to detect the expression in transfected cells, or to simplify the purification process. Suitable fusion partners for the BRCA2 protein or polypeptide are well known in the art and may include  $\beta$ -galactosidase, glutathione-S-transferase, and poly-histidine tag.

Expression vectors may be introduced into host cells by various methods known in the art. The transformation procedure used depends upon the host to be transformed. Methods for introduction of vectors into host cells may include calcium phosphate precipitation, electrosporation, dextran-mediated transfection, liposome encapsulation, nucleus microinjection, and viral or phage infection, among others.

Once an expression vector has been introduced into a suitable host cell, the host cell may be cultured under conditions permitting expression of large amounts of the BRCA2 protein or polypeptide. The expression product may be identified by many approaches well known in the art, for example, sequencing after PCR-based amplification, hybridization using probes complementary to the desired DNA sequence, the presence or absence of marker gene functions such as enzyme activity or antibiotic resistance, the level of mRNA production encoding the intended sequence, immunological detection of a gene product using monoclonal and polyclonal antibodies, such as Western blotting or ELISA. The BRCA2 protein or polypeptides produced in this manner may then be isolated following cell lysis and purified using various protein purification techniques known in the art, for example, ion exchange chromatography, gel filtration chromatography and immunoaffinity chromatography.

It is generally preferred that whenever possible, longer fragments of BRCA2 protein or polypeptide are used, particularly to include the desired functional domains of BRCA2 protein. Expression of shorter fragments of DNA may be useful in generating BRCA2 derived immunogen for the production of anti-BRCA2 antibodies. It should, of course, be understood that not all expression vectors, DNA regulatory sequences or host cells will function equally well to express the BRCA2 protein or polypeptides of the present invention. However, one of ordinary skill in the art may make a selection among expression vectors, DNA regulatory

sequences, host cells, and codon usage in order to optimize expression using known technology in the art without undue experimentation. Studies of BRCA2 protein function and examples of genetic manipulation of BRCA2 protein are summarized in two recent review articles, Bertwistle and Ashworth, *Curr. Opin. Genet. Dev.* 8(1): 14-20 (1998) and Zhang *et al.*, *Cell* 92:433-436 (1998).

### IN VITRO SYNTHESIS AND CHEMICAL SYNTHESIS

Although it is preferred that fragments of the BRCA2 protein or polypeptides be obtained by overexpression in prokaryotic or eukaryotic host cells, the BRCA2 polypeptides or their functional equivalents may also be obtained by *in vitro* translation or synthetic means by methods known to those of ordinary skill in the art. For example, *in vitro* translation may employ an mRNA encoded by a DNA sequence coding for fragments of the BRCA2 protein or polypeptides. Chemical synthesis methodology such as solid phase synthesis may be used to synthesize a BRCA2 polypeptide structural mimic and chemically modified analogs thereof. The polypeptides or the modifications and mimic thereof produced in this manner may then be isolated and purified using various purification techniques, such as chromatographic procedures including ion exchange chromatography, gel filtration chromatography and immunoaffinity chromatography.

### PROTEIN REPLACEMENT THERAPY

The tumor suppressing function of BRCA2 suggests that various BRCA2 protein targeted therapies may be utilized in treating and preventing tumors in breast and ovarian cancer. The present invention therefore includes therapeutic and prophylactic treatment of breast and ovarian cancer using therapeutic pharmaceutical compositions containing the BRCA2 protein, polypeptides, or their functional equivalents. For example, protein replacement therapy may involve directly administering the BRCA2 protein, a BRCA2 polypeptide, or a functional equivalent in a pharmaceutically effective carrier. Alternatively, protein replacement therapy may utilize tumor antigen specific antibody fused to fragments of the BRCA2 protein, a polypeptide, or a functional equivalent to deliver anti-cancer regiments specifically to the tumor cells.

To prepare the pharmaceutical compositions of the present invention, an active BRCA2 protein, a BRCA2 polypeptide, or its functional equivalent is combined with a pharmaceutical carrier selected and prepared according to conventional pharmaceutical compounding techniques. A suitable amount of the composition may be administered locally to the site of a tumor or systemically to arrest the proliferation of tumor cells. The methods for administration, may include parenteral, oral, or intravenous, among others according to established protocols in the art.

Pharmaceutically acceptable solid or liquid carriers or components which may be added to enhance or stabilize the composition, or to facilitate preparation of the composition include, without limitation, syrup, water, isotonic solution, 5 % glucose in water or buffered sodium or ammonium acetate solution, oils, glycerin, alcohols, flavoring agents, preservatives, coloring agents, starches, sugars, diluents, granulating agents, lubricants, binders, and sustained release materials. The dosage at which the therapeutic compositions are administered may vary within a wide range and depends on various factors, such as the stage of cancer progression, the age and condition of the patient, and may be individually adjusted.

### DIAGNOSTIC REAGENTS

The BRCA2 protein, polypeptides, their functional equivalents, antibodies, and polynucleotides may be used in a wide variety of ways in addition to gene therapy and protein replacement therapy. They may be useful as diagnostic reagents to measure normal or abnormal activity of BRCA2 at the DNA, RNA, and protein level. The present invention therefore encompasses the diagnostic reagents derived from the BRCA2 cDNA and protein sequences as set forth in SEQ. ID. Nos: 4-13. These reagents may be utilized in methods for monitoring disease progression, for determining patients suited for gene and protein replacement therapy, or for detecting the presence or quantifying the amount of a tumor growth inhibitor following such therapy. Such methods may involve conventional histochemical techniques, such as obtaining a tumor tissue from the patient, preparing an extract and testing this extract for tumor growth or metabolism. For example, the test for tumor growth may involve measuring abnormal BRCA2 activity using conventional diagnostic assays, such as Southern, Northern, and Western blotting, PCR, RT-PCR, and immunoprecipitation. In

biopsies of tumor tissues, the loss of BRCA2 expression in tumor tissue may be verified by RT-PCR and Northern blotting at the RNA level. A Southern blot analysis, genomic PCR, or fluorescence in situ hybridization (FISH) may also be performed to examine the mutations of BRCA2 at the DNA level. And, a Western blotting, protein truncation assay, or immunoprecipitation may be utilized to analysis the effect at the protein level.

These diagnostic reagents are typically either covalently or non convalently attached to a detectable label. Such a label includes a radioactive label, a colorimetric enzyme label, a fluorescence label, or an epitope label. Frequently, a reporter gene downstream of the regulatory sequences is fused with the BRCA2 protein or polypeptide to facilitate the detection and purification of the target species. Commonly used reporter genes in BRCA2 fusion proteins include  $\beta$ -galactosidase and luciferase gene.

The BRCA2 protein, polypeptides, their functional equivalents, antibodies, and polynucleotides may also be useful in the study of the characteristics of BRCA2 proteins, such as structure and function of BRCA2 in oncogenesis or subcellular localization of BRCA2 protein in normal and cancerous cell. For example, yeast two-hybrid system has been used in the study of cellular function of BRCA2 to identify the regulator and effector of BRCA2 tumor suppressing function (Sharan et al., Nature 386:804-810 (1997) and Katagiri et al., Genes, Chromosomes & Cancer 21:217-222 (1988)). In addition, the BRCA2 protein, polypeptides, their functional equivalents, antibodies, and polynucleotides may also be used in *in vivo* cell based and *in vitro* cell free assays to screen natural products and synthetic compounds which may mimic, regulate or stimulate BRCA2 protein function.

### ANTISENSE INHIBITION

Antisense suppression of endogenous BRCA2 expression may assess the effect of BRCA2 protein on cell growth inhibition using known method in the art (Crooke, *Annu. Rev. Pharmacol. Toxicol.* 32:329-376 (1992) and Robinson-Benion and Holt, *Methods Enzymol.* 254:363-375 (1995)). Given the cDNA sequence as set forth in SEQ ID. NO: 4, 6, 8, 10, and 12, one of skill in the art can readily obtain anti-sense strand of DNA and RNA sequences to interfere with the production of wild-type BRCA2 protein or the mutated form of BRCA2 protein. Alternatively,

antisense oligonucleotide may be designed to target the control sequences of BRCA2 gene to reduce or prevent the expression of the endogenous BRCA2 gene.

### **ANTIBODIES**

The BRCA2 protein, polypeptides, or their functional equivalents may be used as immunogens to prepare polyclonal or monoclonal antibodies capable of binding the BRCA2 derived antigens in a known manner (Harlow & Lane, Antibodies: A Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, 1988). These antibodies may be used for the detection of the BRCA2 protein, polypeptides, or a functional equivalent in an immunoassay, such as ELISA, Western blot, radioimmunoassay, enzyme immunoassay, and immunocytochemistry. Typically, an anti-BRCA2 antibody is in solution or is attached to a solid surface such as a plate, a particle, a bead, or a tube. The antibody is allowed to contact a biological sample or a blot suspected of containing the BRCA2 protein or polypeptide to form a primary immunocomplex. After sufficient incubation period, the primary immunocomplex is washed to remove any non-specifically bound species. The amount of specifically bound BRCA2 protein or polypeptide may be determined using the detection of an attached label or a marker, such as a radioactive, a fluorescent, or an enzymatic label. Alternatively, the detection of BRCA2 derived antigen is allowed by forming a secondary immunocomplex using a second antibody which is attached with a such label or marker. The antibodies may also be used in affinity chromatography for isolating or purifying the BRCA2 protein, polypeptides or their functional equivalents.

### **EXAMPLE 1**

### <u>Determination of the Coding Sequence Haplotypes of the BRCA2 Gene From Normal Individuals</u>

Approximately 150 volunteers were screened in order to identify individuals with no cancer history in their immediate family (i.e. first and second degree relatives). Each person was asked to fill out a hereditary cancer prescreening questionnaire (See TABLE I). Five of these were randomly chosen for end-to-end sequencing of their BRCA2 gene. A first degree relative is a parent, sibling, or

offspring. A second degree relative is an aunt, uncle, grandparent, grandchild, niece, nephew, or half-sibling.

Genomic DNA was isolated from white blood cells of five normal subjects selected from analysis of their answers to the questions above. Dideoxy sequence analysis was performed following polymerase chain reaction amplification.

All exons of the BRCA2 gene were subjected to direct dideoxy sequence analysis by asymmetric amplification using the polymerase chain reaction (PCR) to generate a single stranded product amplified from this DNA sample. Shuldiner, *et al.*, *Handbook of Techniques in Endocrine Research*, p. 457-486, DePablo, F., Scanes, C., eds., Academic Press, Inc., 1993. Fluorescent dye was attached for automated sequencing using the Taq Dye Terminator Kit (Perkin-Elmer<sup>®</sup> cat# 401628). DNA sequencing was performed in both forward and reverse directions on an Applied Biosystems, Inc. (ABI) automated sequencer (Model 377). The software used for analysis of the resulting data was "Sequence Navigator" purchased through ABI.

### 1. Polymerase Chain Reaction (PCR) Amplification

Genomic DNA (100 nanograms) extracted from white blood cells of five normal subjects. Each of the five samples was sequenced end to end. Each sample was amplified in a final volume of 25 microliters containing 1 microliter (100 nanograms) genomic DNA, 2.5 microliters 10X PCR buffer (100 mM Tris, pH 8.3, 500 mM KCl, 1.2 mM MgCl<sub>2</sub>), 2.5 microliters 10X dNTP mix (2 mM each nucleotide), 2.5 microliters forward primer, 2.5 microliters reverse primer, and 1 microliter Taq polymerase (5 units), and 13 microliters of water.

The primers in TABLE II below were used to carry out amplification of the various sections of the BRCA2 gene samples. The primers were synthesized on an DNA/RNA Synthesizer Model  $394^{\$}$ .

Thirty-five cycles were performed, each consisting of denaturing (95°C; 30 seconds), annealing (55°C; 1 minute), and extension (72°C; 90 seconds), except during the first cycle in which the denaturing time was increased to 5 minutes, and during the last cycle in which the extension time was increased to 5 minutes.

PCR products were purified using Qia-quick<sup>®</sup> PCR purification kits (Qiagen<sup>®</sup>, cat# 28104; Chatsworth, CA). Yield and purity of the PCR product are determined spectrophotometrically at OD<sub>260</sub> on a Beckman DU 650 spectrophotometer.

### 2. Dideoxy Sequence Analysis

Fluorescent dye was attached to PCR products for automated sequencing using the Taq Dye Terminator Kit (Perkin-Elmer<sup>®</sup> cat # 401628). DNA sequencing was performed in both forward and reverse directions on an Applied Biosystems, Inc. (ABI) Foster City, CA., automated sequencer (Model 377). The software used for analysis of the resulting data was "Sequence Navigator<sup>®</sup>" purchased through ABI.

### 3. RESULTS

Based upon the sequencing of the five normal individuals, it was determined that the standard sequence found in both GenBank and BIC were inaccurate. In Genbank, a 10 bp stretch (5'-TTTATTTTAG-3') was mistakenly listed as exonic at the 5' end of exon 5 while it should be intronic which would not be included in the cDNA and resultant protein. In addition, a more detrimental error that has the significant potential to lead to an incorrect diagnosis of breast cancer propensity exists in both Genbank and BIC: a sequence of 16 bp (5'-GTGTTCTCATAAACAG-3') should be at the end of exon 15, but instead is listed at the beginning of exon 16 in the database. The disclosure and listing of GenBank is shown in Figure 1. The correct intron/exon sequence of BRCA2 is presented in Figure 2, wherein,

- (1) a 10 bp stretch (5'-TTTATTTTAG-3') is intronic at 3' end of intron 4, rather than at the 5' end of exon 5 (corrected exon 5 is listed as SEQ. ID. NO: 1) and
- (2) a 16 bp stretch (5'-GTGTTCTCATAAACAG-3') is exonic at the 3' end of exon 15, rather than at the 5' end of exon 16 (corrected exons 15 and 16 are listed as SEQ. ID. No: 2 and 3 respectively)

The BIC BRCA2 sequence also contains sequence errors in which a strech of nine nucleotides at positions 5554-5460 is listed as CGTTTGTGT (amino acids: Arg-

Leu-Cys). The correct sequence at these positions is GTTTGTGTT (amino acids: Val-Cys-Val). In addition, the BIC BRCA2 nuclotides at positions 2024 (codon 599), 4553 (codon 1442), 4815 (codon 1529), 5841 (codon 1871), and 5972 (codon 1915) are T, T, A, C, and T respectively, wherein the correct nucleotides at these positions are C, C, G, T, and C respectively. Among them, the nuclotide errors at codon 599, 1442, 1915 result in amino acids changes.

Additional differences in the nucleic acids of the five normal individuals were found in ten polymorphic locations. The changes and their positions are found in TABLE III. The individual haplotypes of each chromosome of BRCA2 are displayed in FIGURE 3. In each case, the initial haplotype reported in Genbank (accession number U43746) was subtracted to determine the new haplotypes OMI 1-5. Thus, the Genbank sequence only represents 50% of the haplotypes found; the five new BRCA2 (OTTI 1-5) DNA sequences are shown as SEQ. ID. NO: 4, 6, 8, 10, and 12, respectively (See FIGURE 3), and the corresponding polypeptides are listed as SEQ. ID. NO: 5, 7, 9, 11, and 13 respectively. In combination, these seven haplotypes represent a functional allele profile for the BRCA2 gene.

The data show that for each of the samples, all exons of BRCA2 were identical except in the region of ten polymorphisms. Six of these polymorphisms were previously identified (Tartigan et al., Nature Genetics 12: 333-337 (1996); Phelan et al., Nature Genetics 13: 120-122 (1996); Couch et al., Nature Genetics 13: 123-125 (1996); Teng, et al., Nature Genetics 13: 241-244 (1996); Schubert et al. 60: 1031-1040 (1997)), but four were unique to this work. Even though the individual polymorphisms may have been identified, none of these complete haplotypes has been previously determined.

### **TABLE I**

### Hereditary Cancer Pre-Screening Questionnaire

### Part A: Answer the following questions about your family

- 1. To your knowledge, has anyone in your family been diagnosed with a very specific hereditary colon disease called Familial Adenomatous Polyposis (FAP)?
- 2. To your knowledge, have you or any aunt had breast cancer diagnosed before the age 35?
- 3. Have you had Inflammatory Bowel Disease, also called Crohn's Disease or Ulcerative Colitis, for more than 7 years?

### Part B: Refer to the list of cancers below for your responses only to questions in Part B

Bladder Cancer Lung Cancer Pancreatic Cancer
Breast Cancer Gastric Cancer Prostate Cancer
Colon Cancer Malignant Melanoma Renal Cancer
Endometrial Cancer Ovarian Cancer Thyroid Cancer

- 4. Have your mother or father, your sisters or brothers or your children had any of the listed cancers?
- Have there been diagnosed in your <u>mother</u>'s brothers or sisters, or your <u>mother</u>'s parents more than one of the cancers in the above list?
- 6. Have there been diagnosed in your <u>father</u>'s brothers or sisters, or your <u>father</u>'s parents <u>more than one</u> of the cancers in the above list?

### Part C: Refer to the list of relatives below for responses only to questions in Part C

You Your mother

Your sisters or brothers Your mother's sisters or brothers (maternal aunts &

uncles)

Your children Your mother's parents (maternal grandparents)

- 7. Have there been diagnosed in these relatives <u>2 or more identical</u> types of cancer? Do not count "simple" skin cancer, also called basal cell or squamous cell skin cancer.
- 8. Is there a <u>total of 4 or more</u> of any cancers in the list of relatives above other than "simple" skin cancers?

### Part D: Refer to the list of relatives below for responses only to questions in Part D.

You Your father

Your sisters or brothers Your father's sisters or brothers (paternal aunts and

uncles)

Your children Your father's parents (paternal grandparents)

- 9. Have there been diagnosed in these relatives <u>2 or more identical</u> types of cancer? Do not count "simple" skin cancer, also called basal cell or squamous cell skin cancer.
- 10. Is there a total of 4 or more of any cancers in the list of relatives above other than "simple" skin cancers?

© Copyright 1996, OncorMed, Inc.

Label	SEQUENCE (5' TO 3') NOTE: M13 TAIL INCLUDED M13 FORWARD = TGT AAA ACG ACG GCC AGT M13 REVERSE = CAG GAA ACA GCT ATG ACC	Oligo Length	PCR Product Length	SEQ. ID. Number
1	5'-TGA GTT TTA CCT CAG TCA CA-3'	20	263	14
	5'-CAG GAA ACA GCT ATG ACC CTG TGA CGT ACT GGG TTT TTA GC-3'	41		15
1	5'-GAT CTT TAA CTG TTC TGG GTC ACA-3'	24	364	16
	5'-CCC AGC ATG ACA CAA TTA ATG A-3'	22		17
1	5'-TGT AAA ACG ACG GCC AGT AGA ATG CAA ATT TAT AAT CCA GAG TA-3'	44	268	18
	5'-ATC AGA TTC ATC TTT ATA GAA C-3'	22		19
BRCA2-5+6F/M13F	5'-TGT AAA ACG ACG GCC AGT TGT GTT GGC ATT TTA AAC ATC A-3'	40	453	20
BRCA2-5+6R/M13R	5'-CAG GAA ACA GCT ATG ACC CAG GGC AAA GGT ATA ACG CT-3'	38		21
	5'-TGT AAA ACG ACG GCC AGT TAA GTG AAA TAA AGA GTG AA-3'	38	248	22
1	5'-CAG GAA ACA GCT ATG ACC AGA AGT ATT AGA GAT GAC-3'	36		23
	5'-TGT AAA ACG ACG GCC AGT GCC ATA TCT TAC CAC CTT GTG A-3'	40	319	24
	5'-TTG CAT TCT AGT GAT AAT ATA C-3'	22	143	25
1	5'-AAT TGT TAG CAA TTT CAA C-3'	19		26
	5'-TGT AAA ACG ACG GCC AGT TGG ACC TAG GTT GAT TGC AGA T-3'	40	338	27
1	5'-CAG GAA ACA GCT ATG ACC TAA ACT GAG ATC ACG GGT GAC A-3'	40		28
1	5'-GAA TAA TAT AAA TTA TAT GGC TTA-3'	24	255	29
BRCA2-10AR/M13R	5'-CAG GAA ACA GCT ATG ACC CCT AGT CTT GCT AGT TCT T-3'	37		30
BRCA2-10BF/M13F	5'- TGT AAA ACG ACG GCC AGT ARC TGA AGT GGA ACC AAA TGA TAC-3'	42	621	31
BRCA2-10BR/M13R	5'- CAG GAA ACA GCT ATG ACC ACG TGG CAA AGA ATT CTC TGA AGT AA-3'	44		32
١				

H S		SEQUENCE (5' TO 3') NOTE: M13 TAIL INCLUDED	Oligo	PCR	SEQ. ID.
		M13 FORWARD = TGT AAA ACG ACG GCC AGT M13 REVERSE = CAG GAA ACA GCT ATG ACC	Length	Product Length	Number
10C	BRCA2-10CF/M13F	5'-TGT AAA ACG ACG GCC AGT CAG CAT CTT GAA TCT CAT ACA G-3'	40	208	33
10C	BRCA2-10CRII	5'-AGA CAG AGG TAC CTG AAT C-3'	19		34
-	BRCA2-11AF-M13	5'- TGT AAA ACG ACG GCC AGT TGG TAC TTT AAT TTT GTC ACT T-3'	40	304	35
11	BRCA2-11AR-M13	5'-CAG GAA ACA GCT ATG ACC TGC AGG CAT GAC AGA GAA T-3'	37		36
11	BRCA2-11BF	5'-AAG AAG CAA AAT GTA ATA AGG A-3'	22	411	37
11	BRCA2-11BR	5'-CAT TTA AAG CAC ATA CAT CTT G-3'	22		38
11	BRCA2-11CF	5'-TCT AGA GGC AAA GAA TCA TAC-3'	21	349	39
11	BRCA2-11CR	5'-CAA GAT TAT TCC TTT CAT TAG C-3'	22		40
11	BRCA2-11DF	5'-AAC CAA AAC ACA AAT CTA AGA G-3'	22	344	41
11	BRCA2-11DR	5'-GTC ATT TTT ATA TGC TGC TTT AC-3'	23		42
1	BRCA2-11EF	5'-GGT TTT ATA TGG AGA CAC AGG-3'	21	369	43
17	BRCA2-11ER	5'-GTA TTT ACA ATT TCA ACA CAA GC-3'	23		44
7	BRCA2-11FF	5'-ATC ACA GTT TTG GAG GTA GC-3'	20	368	45
=	BRCA2-11FR	5'-CTG ACT TCC TGA TTC TTC TAA-3'	21		46
17	BRCA2-11GF	5'-CTC AGA TGT TAT TTT CCA AGC-3'	21	366	47
1-	BRCA2-11GR	5'-CTG TTA AAT AAC CAG AAG CAC-3'	21		48
11	BRCA2-11HF	5'-AGG TAG ACA GCA AGC-3'	18	360	49
=	BRCA2-11HR	5'-GTA ATA TCA GTT GGC ATT TAT T-3'	22		50
11	BRCA2-11IF	5'-TGC AGA GGT ACA TCC AAT AAG-3'	21	326	51
11	BRCA2-11IR	5'-GAT CAG TAA ATA GCA AGT CCG-3'	21		52
11	BRCA2-11JF	5'-TAC TGA AAA TGA AGA TAA CAA AT-3'	23	477	53

	ED         Oligo         PCR         SEQ. ID.           ACG GCC AGT         Length         Product         Number           GCT ATG ACC         Length         Length	22	CGG AGC AA-3' 35 382 55	CCC AAC AG-3' 35 56	22 374 57	19 58	20 409 59	22 60	35 306 61	GTA GGA AT-3' 35 62	22 383 63	20 64	20 355 65	20 66	21 337 67	21 68	20 360 69	20 70	TGG AAA GC-3' 35 458 71	TTT TAC CAA T-3' 37 72	22 344 73	21 74
SEQUENCE (5' TO 3') NOTE:	M13 FORWARD = TGT AAA ACG ACG GCC AGT M13 FORWARD = TGT AAA ACG ACG GCC AGT M13 REVERSE = CAG GAA ACA GCT ATG ACC	5'-ATT TTG TTC TTT CTT ATG TCA G-3'	5'-TGT AAA ACG ACG GCC AGT CTA CTA AAA CGG AGC AA-3'	5'-CAG GAA ACA GCT ATG ACC GTA TGA AAA CCC AAC AG-3'	5-CAC AAA ATA CTG AAA GAA AGT G-3'	5'-GGC ACC ACA GTC TCA ATA G-3'	5'-GCA AAG ACC CTA AAG TAC AG-3'	5'-CAT CAA ATA TTC CTT CTC TAA G-3'	5'-TGT AAA ACG ACG GCC AGT GAA AAT TCA GCC TTA GC-3'	5'- CAG GAA ACA GCT ATG ACC ATC AGA ATG GTA GGA AT-3'	5'-GTA CTA TAG CTG AAA ATG ACA A-3'	5'-ACC ACT GGC TAT CCT AAA TG-3'	5'-TGA AGA TAT TTG CGT TGA GG-3'	5'-GTC AGC AAA AAC CTT ATG TG-3'	5'-ACG AAA ATT ATG GCA GGT TGT-3'	5'-CTT GTC TTG CGT TTT GTA ATG-3'	5'-GCT TCA TAA GTC AGT CTC AT-3'	5'-TCA AAT TCC TCT AAC ACT CC-3'	5'-TGT AAA ACG ACG GCC AGT TAC AGC AAG TGG AAA GC-3'	5'-CAG GAA ACA GCT ATG ACC AAG TTT CAG TTT TAC CAA T-3'	5'-GTT CTT CAG AAA ATA ATC ACT C-3'	5'-TGT AAA AAG AGA ATG TGT GGC-3'
	Label	BRCA2-11JR 5	BRCA2-11KF-M13 5	BRCA2-11KR-M13 5	BRCA2-11LF	BRCA2-11LR 5	BRCA2-11MF	BRCA2-11MR	BRCA2-11NF-M13 5	BRCA2-11NR-M13 E	BRCA2-110F	BRCA2-110R	BRCA2-11PF	BRCA2-11PR	BRCA2-11QF	BRCA2-11QR	BRCA2-11RF	BRCA2-11RR	BRCA2-11SF-M13	BRCA2-11SR-M13	BRCA2-11TF	BRCA2-11TR
	Exon	11	11	11	11	11	11	11	11	1.1	1.		11	11	11	-	11	11	11	11	11	11

2	-	SEQUENCE (5' TO 3') NOTE:	Oligo	PCR	SEO. ID.
		M13 FORWARD = TGT AAA ACG ACG GCC AGT M13 REVERSE = CAG GAA ACA GCT ATG ACC	Length	Product Length	Number
11	BRCA2-11UF-M13	5'-TGT AAA ACG ACG GCC AGT ACT TTT TCT GAT GTT CCT GTG-3'	39	328	52
11	BRCA2-11UR-M13	5'-CAG GAA ACA GCT ATG ACC TAA AAA TAG TGA TTG GCA ACA-3'	39		92
12	BRCA2-12F/M13F	5'-TGT AAA ACG ACG GCC AGT AGT GGT GTT TTA AAG TGG TCA AAA-3'	42	391	77
12	BRCA2-12R/M13R	5'-CAG GAA ACA GCT ATG ACC GGA TCC ACC TGA GGT CAG AAT A-3'	40		78
13	BRCA2/13-2F	5'-TAA CAT TTA AGC ATC CGT TAC-3'	21	310	79
13	BRCA2/13-2R	5'-AAA CGA GAC TTT TCT CAT ACT GTA TTA G-3'	28		80
14	BRCA2-14F	5'-ACC ATG TAG CAA ATG AGG GTC T-3'	22	391	81
14	BRCA2-14AR	5'-GCT TTT GTC TGT TTT CCT CCA A-3'	22		82
15	BRCA2-15-2F	5'-CCA GGG GTT GTG CTT TTT AAA-3'	21	284	83
15	BRCA2-15FUT/M13-R	5'-CAG GAA ACA GCT ATG ACC ACT CTG TCA TAA AAG CCA TC-3'	38		84
16	BRCA2-16AF	5'-TTT GGT TTG TTA TAA TTG TTT TTA-3'	24	394	85
16	BRCA2-16AR	5'-CCA ACT TTT TAG TTC GAG AG-3'	50		86
17	BRCA2-17F	5'-TTC AGT ATC ATC CTA TGT G-3'	19	282	87
17	BRCA2-17AR	5'-AGA AAC CTT AAC CCA TAC TG-3'	20		88
18	BRCA2-18FUT/M13-AF	BRCA2-18FUT/M13-AF 5'-TGT AAA ACG ACG GCC AGT GAA TTC TAG AGT CAC ACT TCC-3'	39	275	89
18	BRCA2-18R/M13R	5'-CAG GAA ACA GCT ATG ACC TTT AAC TGA ATC AAT GAC TG-3'	38		06
19	BRCA2-19F/M13F	5'-TGT AAA ACG ACG GCC AGT AAG TGA ATA TTT TTA AGG CAG TT-3'	41	355	91
19	BRCA2-19FUT/M13-R	5'-CAG GAA ACA GCT ATG ACC AAG AGA CCG AAA CTC CAT CTC-3'	39		92
20	BRCA2-20F/M13F	5'-TGT AAA ACG ACG GCC AGT CAC TGT GCC TGG CCT GAT AC-3'	38	296	93
20	BRCA2-20R/M13R	5'-CAG GAA ACA GCT ATG ACC ATG TTA AAT TCA AAG TCT CTA-3'	36		94
21	BRCA2-21F/M13F	5'-TGT AAA ACG ACG GCC AGT GGG TGT TTT ATG CTT GGT TCT-3'	39	304	95

Exon	Label	SEQUENCE (5' TO 3') NOTE: M13 TAIL INCLUDED M13 FORWARD = TGT AAA ACG ACG GCC AGT M13 REVERSE = CAG GAA ACA GCT ATG ACC	Oligo Length	PCR Product Length	SEQ. ID. Number
21	BRCA2-21R/M13R	5-CAG GAA ACA GCT ATG ACC CAT TTC AAC ATA TTC CTT CCT G-3'	40		96
22	BRCA2-22F-1A	5'-AAC CAC ACC CTT AAG ATG A-3'	19	453	97
22	BRCA2-22R-1A	5'-GCA TTA GTA GTG GAT TTT GC-3'	20		98
23	BRCA2-23FII	5'-TCA CTT CCA TTG CAT C-3'	16	290	66
23	BRCA2-23RII	5'-TGC CAA CTG GTA GCT CC-3'	17		100
24	BRCA2-24 2F	5'-TAC AGT TAG CAG CGA CAA AA-3'	20	373	101
24	BRCA2-24R/M13R	5'-CAG GAA ACA GCT ATG ACC ATT TGC CAA CTG GTA GCT CC-3'	38		102
25	BRCA2-25F-7/23	5'-GCT TTC GCC AAA TTC AGC TA-3'	20	427	103
25	BRCA2-25R-7/23	5'-TAC CAA AAT GTG TGG TGA TG-3'	20		104
26	BRCA2/26-2F	5'-AAT CAC TGA TAC TGG TTT TG-3'	20	530	105
26	BRCA2/26-2R	5'-TAT ACT TAC AGG AGC CAC AT-3'	20		106
27A	BRCA2-27AF-1A	5'-CTG TGT GTA ATA TTT GCG-3'	18	495	107
27A	BRCA2-27AR/M13R	5'-CAG GAA ACA GCT ATG ACG GCA AGT TCT TCG TCA GCT ATT G-3'	40		108
278	BRCA2-27BF/M13F	5'-TGT AAA ACG ACG GCC AGT GAA TTC TCC TCA GAT GAC TCC A-3'	40	417	109
27B	BRCA2-27BR/M13R	5'-CAG GAA ACA GCT ATG ACC TCT TTG CTC ATT GTG CAA CA-3'	38		110

TABLE III NORMAL PANEL TYPING

Position nt/codon	Nucleotide Change	Amino Acid Change	_	2	3	4	5	Frequency
1093/289	AAT → CAT	Asn → His	A/A	AC	A/A	A/A	A/C	A = .8 C = .2
1342/372	AAT → CAT	Asn → His	AC	A/A	A/C	A/C	A/C	A = 0.6 C = 0.4
1593/455	TC <u>A</u> → TC <u>G</u>	Ser → Ser	ΝΑ	A/A	A/A	A/A	A/G	A = 0.9 G = 0.1
2457/743	CA <u>I</u> → CA <u>C</u>	His → His	T/T	С/Т	1/1	T/T	СЛ	T = 0.8 C = 0.2
2908/894	GTA → <u>A</u> TA	Val → Ile	9/9	9/9	9/9	9/9	A/G	G = 0.9 A = 0.1
3199/991	<u>A</u> AC → <u>G</u> AC	Asn → Asp	A/A	A/G	A/A	A/A	A/G	A = 0.8 G = 0.2

TABLE III NORMAL PANEL TYPING

Position nt/codon	Nucleotide Change	Amino Acid Change	-	7	က	4	5	Frequency
3624/1132	AA <u>A</u> → AA <u>G</u>	Lys → Lys	A/A	A/G	A/A	A/G	A/A	A = 0.8 G = 0.2
4035/1269	GT <u>I</u> → GT <u>C</u>	Val→Val	C/T	Т/Т	Т/Т	Т/Т	Т/Т	T = 0.9 C = 0.1
7470/2414	TC <u>A</u> → TC <u>G</u>	Ser → Ser	Α/A	A/G	A/A	A/G	A/A	A = 0.8 G = 0.2
9079/2951	GCC → ACC	Ala → Thr	9/9	9/9	9/9	9/9	A/G	G = 0.9 A = 0.1

#### **EXAMPLE 2**

5

10

15

20

25

30

### <u>Determination Of A Normal Individual Using BRCA2<sup>(OMI 1-5)</sup> and The Ten</u> Polymorphisms For Reference

A person skilled in the art of genetic susceptibility testing will find the present invention useful for:

- a) identifying individuals having a normal BRCA2 gene;
- avoiding misinterpretation of normal polymorphisms found in the normal population.

Sequencing was carried out as in EXAMPLE 1 using a blood sample from the patient in question. However, the BRCA2<sup>(omi1-5)</sup> sequences were used for reference and any polymorphic sites seen in the patient were compared to the nucleic acid sequences listed above for normal codons at each polymorphic site. A normal sample is one which is comparable to the BRCA2<sup>(omi 1-5)</sup> sequences and contains only minor variations which occur at minor polymorphic sites. The allelic variations which occur at each of the polymorphic sites are paired here for reference.

- <u>A</u>AT (Asn) and <u>C</u>AT (His) at position 1093 (codon 289)
- <u>CAT</u> (His) and <u>AAT</u> (Asn) at position 1342 (codon 372)
- TCA (Ser) and TCG (Ser) at position 1593 (codon 455)
- CAT (His) and CAC (His) at position 2457 (codon 743)
- <u>G</u>TA (Val) and <u>A</u>TA (IIe) at position 2908 (codon 894)
- AAC (Asn) and GAC (Asp) at position 3199 (codon 991)
- AAA (Lys) and AAG (Lys) at position 3624 (codon 1132)
- GT<u>T</u> (Val) and GT<u>C</u> (Val) at position 4035 (codon 1269)
- TCA (Ser) and TCG (Ser) at position 7470 (codon 2414)
- GCC (Ala) and ACC (Thr) at position 9079 (codon 2951)

The availability of these polymorphic pairs provides added assurance that one skilled in the art can correctly interpret the polymorphic variations without mistaking a normal variation for a mutation.

All exons of the BRCA2 gene are subjected to direct dideoxy sequence analysis by asymmetric amplification using the polymerase chain reaction (PCR) to generate a single stranded product amplified from this DNA sample. Shuldiner, et

10

15

20

25

30

al., Handbook of Techniques in Endocrine Research, p. 457-486, DePablo, F., Scanes, C., eds., Academic Press, Inc., 1993. Fluorescent dye is attached for automated sequencing using the Taq Dye Terminator Kit (Perkin-Elmer<sup>®</sup> cat# 401628). DNA sequencing is performed in both forward and reverse directions on an Applied Biosystems, Inc. (ABI) automated sequencer (Model 377). The software used for analysis of the resulting data is "Sequence Navigator" purchased through ABI.

#### 1. Polymerase Chain Reaction (PCR) Amplification

The PCR primers used to amplify a patient's sample BRCA2 gene are listed in TABLE II. The primers were synthesized on a DNA/RNA Synthesizer Model 394<sup>®</sup>. Thirty-five cycles are of amplification are performed, each consisting of denaturing (95°C; 30 seconds), annealing (55°C; 1 minute), and extension (72°C; 90 seconds), except during the first cycle in which the denaturing time is increased to 5 minutes and during the last cycle in which the extension time is increased to 5 minutes.

PCR products are purified using Qia-quick<sup>®</sup> PCR purification kits (Qiagen<sup>®</sup>, cat# 28104; Chatsworth, CA). Yield and purity of the PCR product are determined spectrophotometrically at OD<sub>260</sub> on a Beckman DU 650 spectrophotometer.

#### 2. Dideoxy Sequence Analysis

Fluorescent dye is attached to PCR products for automated sequencing using the Taq Dye Terminator Kit (Perkin-Elmer<sup>®</sup> cat# 401628). DNA sequencing is performed in both forward and reverse directions on an Applied Biosystems, Inc. (ABI) Foster City, CA., automated sequencer (Model 377). The software used for analysis of the resulting data is "Sequence Navigator<sup>®</sup>" purchased through ABI. The BRCA2<sup>(omi 1-5)</sup> sequences were entered sequentially into the Sequence Navigator software as the standards for comparison. The Sequence Navigator software compares the patient sample sequence to each BRCA2 <sup>(omi 1-5)</sup> standard, base by base. The Sequence Navigator highlights all differences between the standards (omi 1-5) and the patient's sample sequence.

A first technologist checks the computerized results by comparing visually the BRCA2 (orm 1-5) standards against the patient's sample, and again highlights any differences between the standard and the sample. The first primary technologist then interprets the sequence variations at each position along the sequence. Chromatograms from each sequence variation are generated by the Sequence Navigator and printed on a color printer. The peaks are interpreted by the first primary technologist and a second primary technologist. A secondary technologist then reviews the chromatograms. The results are finally interpreted by a geneticist. In each instance, a variation is compared to known normal polymorphisms for position and base change.

#### 3. Results

5

10

15

20

25

30

The patient's BRCA2 sequence was found to be heterozygous at seven nucleotide positions: 1093 (A/C), 1342 (A/C), 1593 (A/G), 2457 (C/T), 2908 (A/G), 3199 (A/G) and 9079 (A/G). In addition, this changes five amino acids in the polypeptide product: Asn to His at codon 289, Asn to His at codon 372, Val to Ile at codon 894, Asn to Asp at codon 991, and Ala to Thr at codon 2951. The question arises whether any or all of these changes have significance to the patient. Comparison of the patient's results to the BRCA (omi 1-5) haplotypes demonstrates that it matches one of the BRCA2 omi standards (#5), and thus the patient sample is interpreted as carrying a normal gene sequence without causing any elevation in their risk status for breast cancer.

#### **EXAMPLE 3**

# DETERMINING THE PRESENCE OF A MUTATION IN EXON 11 OF THE BRCA2 GENE USING BRCA2<sup>(omi1-5)</sup>

A person skilled in the art of genetic susceptibility testing will find the present invention useful for determining the presence of a known or previously unknown mutation in the BRCA2 gene. A list of mutations of BRCA2 is publicly available in the Breast Cancer Information Core at http://www.nchgr.nih.gov/dir/lab\_transfer/bic. This data site became publicly available on November 1, 1995. Friend, S. *et al. Nature Genetics* 11:238, (1995).

In this example, a mutation in exon 11 is characterized by amplifying the region of the mutation with a primer set which amplifies the region of the mutation. Sequencing was carried out as in Example 1 using a blood sample from the patient in question. Specifically, exon 11 of the BRCA2 gene is subjected to direct dideoxy sequence analysis by asymmetric amplification using the polymerase chain reaction (PCR) to generate a single stranded product amplified from this DNA sample. Shuldiner, et al., Handbook of Techniques in Endocrine Research, p. 457-486, DePablo, F., Scanes, C., eds., Academic Press, Inc., 1993. Fluorescent dye is attached for automated sequencing using the Taq Dye Terminator Kit (Perkin-Elmer<sup>®</sup> cat# 401628). DNA sequencing is performed in both forward and reverse directions on an Applied Biosystems, Inc. (ABI) automated sequencer (Model 377). The software used for analysis of the resulting data is "Sequence Navigator" purchased through ABI.

15

20

25

10

5

#### 1. Polymerase Chain Reaction (PCR) Amplification

Genomic DNA (100 nanograms) extracted from white blood cells of the subject is amplified in a final volume of 25 microliters containing 1 microliter (100 nanograms) genomic DNA, 2.5 microliters 10X PCR buffer (100 mM Tris, pH 8.3, 500 mM KCl, 1.2 mM MgCl<sub>2</sub>), 2.5 microliters 10X dNTP mix (2 mM each nucleotide), 2.5 microliters forward primer (BRCA2-11Q-F, 10 micromolar solution), 2.5 microliters reverse primer (BRCA2-11Q-R, 10 micromolar solution), and 1 microliter Tag polymerase (5 units), and 13 microliters of water.

The PCR primers used to amplify segment Q of exon 11 (where the mutation 6174delT is found) are as follows:

BRCA2-11Q-F: 5'- ACG' AAA' ATT' ATG' GCA' GGT' TGT-3'

BRCA2-11Q-R: 5'- CTT' GTC' TTG' CGT' TTT' GTA' ATG-3'

30

The primers are synthesized on an DNA/RNA Synthesizer Model 394<sup>®</sup>. Thirty-five cycles are performed, each consisting of denaturing (95°C; 30 seconds), annealing (55°C; 1 minute), and extension (72°C; 90 seconds), except during the

10

15

20

25

first cycle in which the denaturing time is increased to 5 minutes, and during the last cycle in which the extension time is increased to 5 minutes.

PCR products are purified using Qia-quick<sup>®</sup> PCR purification kits (Qiagen<sup>®</sup>, cat# 28104; Chatsworth, CA). Yield and purity of the PCR product are determined spectrophotometrically at OD<sub>260</sub> on a Beckman DU 650 spectrophotometer.

#### 2. <u>Dideoxy Sequence Analysis</u>

Fluorescent dye is attached to PCR products for automated sequencing using the Taq Dye Terminator Kit (Perkin-Elmer<sup>®</sup> cat# 401628). DNA sequencing is performed in both forward and reverse directions on an Applied Biosystems, Inc. (ABI) Foster City, CA., automated sequencer (Model 377). The software used for analysis of the resulting data is "Sequence Navigator<sup>®</sup>" purchased through ABI. The BRCA2<sup>(omi 1-5)</sup> sequence is entered into the Sequence Navigator software as the Standard for comparison. The Sequence Navigator software compares the sample sequence to the BRCA2<sup>(omi)</sup> standard, base by base. The Sequence Navigator highlights all differences between the BRCA2<sup>(omi)</sup> normal DNA sequence and the patient's sample sequence.

A first technologist checks the computerized results by comparing visually the BRCA2<sup>(omi 1-5)</sup> standard against the patient's sample, and again highlights any differences between the standard and the sample. The first primary technologist then interprets the sequence variations at each position along the sequence. Chromatograms from each sequence variation are generated by the Sequence Navigator and printed on a color printer. The peaks are interpreted by the first primary technologist and a second primary technologist. A secondary technologist then reviews the chromatograms. The results are finally interpreted by a geneticist. In each instance, a sequence variation is compared to known normal polymorphisms for position and base change. The ten frequent polymorphisms which occur in BRCA2 are:

- AAT (Asn) and CAT (His) at position 1093 (codon 289)
- <u>CAT</u> (His) and <u>AAT</u> (Asn) at position 1342 (codon 372)
- TCA (Ser) and TCG (Ser) at position 1593 (codon 455)

- CAT (His) and CAC (His) at position 2457 (codon 743)
- GTA (Val) and ATA (Ile) at position 2908 (codon 894)
- AAC (Asn) and GAC (Asp) at position 3199 (codon 991)
- AAA (Lys) and AAG (Lys) at position 3624 (codon 1132)
- GTT (Val) and GTC (Val) at position 4035 (codon 1269)
- TCA (Ser) and TCG (Ser) at position 7470 (codon 2414)
- GCC (Ala) and ACC (Thr) at position 9079 (codon 2951)

#### 10 3. Results

5

15

20

25

30

Using the above PCR amplification and standard fluorescent sequencing technology, the 6174delT mutation may be found. Mutations are noted by the length of non-matching sequence variation. Such a lengthy mismatch pattern occurs with deletions and insertions. This mutation is named in accordance with the suggested nomenclature for naming mutations, Beaudet, A *et al.*, *Human Mutation* 2:245-248, (1993). The 6174delT mutation at codon 1982 of the BRCA2 gene lies in segment "Q" of exon 11. The DNA sequence results demonstrate the presence of a one base pair deletion of a T at nucleotide 6174 of the BRCA2 transcript, resulting in the appearance of an in-frame terminator (TAG) at codon position 2003. This mutation is, therefore, predicted to result in a truncated, and most likely, non-functional protein.

#### **EXAMPLE 4**

## GENERATION OF MONOCLONAL AND POLYCLONAL ANTIBODIES USING GST-BRCA2 FUSION PROTEIN AS AN IMMUNOGEN

DNA primers are used to amplify a fragment of BRCA2 using PCR technology. The product is then digested with suitable restriction enzymes and fused in frame with the gene encoding glutathione S-transferase (GST) in *Escherichia coli* using GST expression vector pGEX (Pharmacia Biotech Inc.) The expression of the fusion protein is induced by the addition of isopropyl-β-thiogalactopyranoside. The bacteria are then lysed and the overexpressed fusion protein is purified with glutathione-sepharose beads. The fusion protein is then verified by SDS/PAGE gel and N-terminus protein sequencing. The purified protein

10

15

20

25

30

is used to immunize rabbits according to standard procedures described in Harlow & Lane (1988). Polycolonal antibody is collected from the serum several weeks after and purified using known methods in the art. Monoclonal antibodies against all or fragments of BRCA2 protein, polypeptides, or functional equivalents are obtained using hybridoma technology, see also Harlow & Lane (1988). The BRCA2 protein or polypeptide is coupled to the carrier keyhole limpet hemocyanin in the presence of glutaraldehyde. The conjugated immunogen is mixed with an adjuvant and injected into rabbits. Spleens from antibody-containing rabbits are removed. The B-cells isolated from spleen are fused to myeloma cells using polyethylene glycol (PEG) to promote fusion. The hybrids between the myeloma and B-cells are selected and screened for the production of antibodies to immunogen BRCA2 protein or polypeptide. Positive cells are recloned to generate monoclonal antibodies.

#### **EXAMPLE 5**

#### DETECTION OF BRCA2 EXPRESSION IN HUMAN TISSUES AND CELL LINES

The expression of BRCA2 in human tissues is determined using Northern blot analysis. Human tissues include those from pancreas, testis, prostate, ovary, breast, small intestine, and colon are obtained from Clontech Laboratories, Inc., Palo Alto, CA. The poly(A)+ mRNA Northern blots from different human tissues is hybridized to BRCA2 cDNA probes according to manufacture protocol. The expression level is further conformed by RT-PCR using oligo-d(T) as a primer and other suitable primers.

For Northern Blot analysis of cancer cell lines, the human ovarian cancer cell line SKOV-3 and the human breast cancer cell line MCF-7 are obtained from the American Type Culture Collection. Total RNA is prepared by lysing cell in the presence of guanidinium isocyanate. Poly(A)\* mRNA is isolated using the PolyATract mRNA isolation system from Promega, Madison, WI. The isolated RNA is then electrophoresed under denaturing conditions and transferred to Nylon membrane. The probe used for Northern blot is a fragment of BRCA2 sequence obtained by PCR amplification. The probes are labeled with [ $\alpha$ -32P] dCTP using a random-primed labeling kit (Amersham Life Science, Arlington Heights, IL).

#### **EXAMPLE 6**

5

10

15

20

25

30

#### **EXPRESSION OF THE BRCA2 PROTEIN**

The whole-cell extracts of BRCA2 transfected cells are subjected to immunoprecipitation and immunoblotting to determine the BRCA2 protein level. The BRCA2 protein or polypeptide is immunoprecipitated using anti-BRCA2 antibodies prepared according to Example 4. Samples are then fractionated using SDS/PAGE gel and transferred to nitrocellulose. Western blot of the BRCA2 protein or polypeptide is performed with the indicated antibodies. Antibody reaction is revealed using enhanced chemiluminescence reagents (Dupont New England Nuclear, Boston, MA).

#### **EXAMPLE 7**

#### USE OF THE BRCA2(omi1-5) GENE THERAPY

The growth of ovarian or breast cancer may be arrested by increasing the expression of the BRCA2 gene where inadequate expression of that gene is responsible for hereditary ovarian or breast cancer. Gene therapy may be performed on a patient to reduce the size of a tumor. The LXSN vector may be transformed with a BRCA2<sup>(omi1-5)</sup> coding sequence as presented SEQ ID NO:4, 6, 8, 10, or 12 or a fragment thereof.

#### Vector

The LXSN vector is transformed with a fragment of the wildtype BRCA2<sup>(omi1-5)</sup> coding sequence as set forth in SEQ ID NO:4, 6, 8, 10, or 12. The LXSN-BRCA2<sup>(omi1-5)</sup> retroviral expression vector is constructed by cloning a *Sal* I linkered BRCA2<sup>(omi1-5)</sup> cDNA or fragments thereof into the *Xho* I site of the vector LXSN. Constructs are confirmed by DNA sequencing. See Holt et al., *Nature Genetics* 12: 298-302 (1996). Retroviral vectors are manufactured from viral producer cells using serum free and phenol-red free conditions and tested for sterility, absence of specific pathogens, and absence of replication-competent retrovirus by standard assays. Retrovirus is stored frozen in aliquots which have been tested.

Patients receive a complete physical exam, blood, and urine tests to determine overall health. They may also have a chest X-ray, electrocardiogram, and appropriate radiologic procedures to assess tumor stage.

Patients with metastatic ovarian cancer are treated with retroviral gene therapy by infusion of recombinant LXSN-BRCA2<sup>(om/1-5)</sup> retroviral vectors into peritoneal sites containing tumor, between 10<sup>9</sup> and 10<sup>10</sup> viral particles per dose. Blood samples are drawn each day and tested for the presence of retroviral vector by sensitive polymerase chain reaction (PCR)-based assays. The fluid which is removed is analyzed to determine:

- 1. The percentage of cancer cells which are taking up the recombinant LXSN-BRCA2<sup>(omi1-5)</sup> retroviral vector combination. Successful transfer of BRCA1 gene into cancer cells has been shown by both RT-PCR analysis and *in situ* hybridization. RT-PCR is performed with by the method of Thompson et al., *Nature Genetics* 9: 444-450 (1995), using primers derived from a BRCA2<sup>(omi1-5)</sup> coding sequence as in SEQ ID NO:4, 6, 8, 10, or 12 or fragments thereof. Cell lysates are prepared and immunoblotting is performed by the method of Jensen *et al.*, *Nature Genetics* 12: 303-308 (1996) and Jensen *et al.*, *Biochemistry* 31: 10887-10892 (1992).
- 2. Presence of programmed cell death using APOTAG® in situ apoptosis detection kit (ONCOR, INC., Gaithersburg, Maryland) and DNA analysis.
- 3. Measurement of BRCA2 gene expression by slide immunofluorescence or Western blot.

Patients with measurable disease are also evaluated for a clinical response to LXSN-BRCA2<sup>(omi1-5)</sup> especially those that do not undergo a palliative intervention immediately after retroviral vector therapy. Fluid cytology, abdominal girth, CT scans of the abdomen, and local symptoms are followed.

For other sites of disease, conventional response criteria are used as follows:

- 1. Complete Response (CR), complete disappearance of all measurable lesions and of all signs and symptoms of disease for at least 4 weeks.
- 2. Partial Response (PR), decrease of at least 50% of the sum of the products of the 2 largest perpendicular diameters of all measurable lesions as determined by 2 observations not less than 4 weeks apart. To be considered a PR, no new lesions should have appeared during this period and none should have increased in size.
- 3. Stable Disease, less than 25% change in tumor volume from previous evaluations.

25

30

5

10

15

15

20

25

4. Progressive Disease, greater than 25% increase in tumor measurements from prior evaluations. The number of doses depends upon the response to treatment.

#### 5 **EXAMPLE 8**

#### PROTEIN REPLACEMENT THERAPY

Therapeutically elevated level of functional BRCA2 protein may alleviate the absence or reduced endogenous BRCA2 tumor suppressing activity. Breast or ovarian cancer is treated by the administration of a therapeutically effective amount of the BRCA2 protein, a polypeptide, or its functional equivalent in a pharmaceutically acceptable carrier. Clinically effective delivery method is applied either locally at the site of the tumor or systemically to reach other metastasized locations with known protocols in the art. These protocols may employ the methods of direct injection into a tumor or diffusion using time release capsule. A therapeutically effective dosage is determined by one of skill in the art.

Breast or ovarian cancer may be prevented by the administration of a prophylactically effective amount of the BRCA2 protein, polypeptide, or its functional equivalent in a pharmaceutically acceptable carrier. Individuals with known risk for breast or ovarian cancer are subjected to protein replacement therapy to prevent tumorigenesis or to decrease the risk of cancer. Elevated risk for breast and ovarian cancer includes factors such as carriers of one or more known BRCA1 and BRCA2 mutations, late child bearing, early onset of menstrual period, late occurrence of menopause, and certain high risk dietary habits. Clinically effective delivery method is used with known protocols in the art, such as administration into peritoneal cavity, or using an implantable time release capsule. A prophylactically effective dosage is determined by one of skill in the art.

#### TABLE OF REFERENCES

- 1. Sanger, F., et al., J. Mol. Biol. 42:1617, (1980).
- 2. Beaucage, et al., Tetrahedron Letters <u>22</u>:1859-1862, (1981).
  - 3. Maniatis, et al. in Molecular Cloning: A Laboratory Manual, Cold Spring Harbor, NY, p 280-281, (1982).
  - 4. Conner, et al., Proc. Natl. Acad. Sci. U.S.A. 80:278, (1983)
  - 5. Saiki, et.al.., *Bio/Technology* 3:1008-1012, (1985)

30

- 6. Landgren, et al., Science 241:1007,(1988)
- 7. Landgren, et al., Science 242:229-237, (1988).
- 8. PCR. A Practical Approach, ILR Press, Eds. M. J. McPherson, P. Quirke, and G. R. Taylor, (1992).
  - 9. Easton et al., American Journal of Human Genetics 52:678-701, (1993).
  - 10. Patent No. 4,458,066.
  - 11. Rowell, S., et al., American Journal of Human Genetics 55:861-865, (1994)
  - 12. Miki, Y. et al., Science 266:66-71, (1994).
- 10 13. Wooster, R. et al., Science <u>265</u>:2088-2090, (1994).
  - 14. Wooster, R. et al., Nature 378:789-792, (1995).
  - 15. Beaudet, A et al., Human Mutation 2:245-248, (1993).
  - 16. Friend, S. et al. Nature Genetics 11:238, (1995).
  - 17. Teng et al, Nature Genetics 13: 241-244 (1996).
- 15 **18. Couch** *et al*, *Nature Genetics* <u>13</u>: 123-125 (1996).
  - 19. Tartigan et al, Nature Genetics 12: 333-337 (1996).
  - 20. Phelan et al, Nature Genetics 13: 120-122 (1996).
  - 21. Schubert et al, American Journal of Human Genetics 60: 1031-1040 (1996).
  - 22. Sambrook, et al., Molecular Cloning: A Laboratory Manual, Second Edition, Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y. (1989).
  - 23. Bertwistle and Ashworth, Curr. Opin. Genet. Dev. 8(1): 14-20 (1998).
  - 24. Zhang et al., Cell 92:433-436 (1998).
  - 25. Sharan et al., Nature 386:804-810 (1997).
  - 26. Katagiri et al., Genes, Chromosomes & Cancer 21:217-222 (1988).
- 25 27. Crooke, Annu. Rev. Pharmacol. Toxicol. 32:329-376 (1992)
  - 28. Robinson-Benion and Holt, Methods Enzymol. 254:363-375 (1995).
  - 29. Harlow & Lane, Antibodies: A Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, 1988.
  - 30. Shuldiner, et al., Handbook of Techniques in Endocrine Research, p. 457-486, DePablo, F., Scanes, C., eds., Academic Press, Inc., 1993.
    - 31. Holt et al., Nature Genetics 12: 298-302 (1996).
    - 32. Thompson et al., Nature Genetics 9: 444-450 (1995).
    - 33. Jensen et al., Nature Genetics 12: 303-308 (1996)
    - 34. Jensen et al., Biochemistry 31: 10887-10892 (1992).

Although the invention has been described with reference to the presently preferred embodiments, it should be understood that various modifications can be made without departing from the spirit of the invention. Accordingly, the invention is limited only by the following claims.

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:(A) LENGTH: 50 base pairs

(1) GENERAL INFORMATION

(i) APPLICANT: Murphy, Patricia

White, Marga

5

60

SEQUENCE LISTING

5	<ul><li>(B) TYPE: nucleic acid</li><li>(C) STRANDEDNESS: double</li><li>(D) TOPOLOGY: linear</li></ul>	
J	<pre>(ii) MOLECULE TYPE: Genomic DNA (ix) FEATURE:</pre>	
10	<ul><li>(A) NAME/KEY: exon</li><li>(B) LOCATION: 150</li><li>(D) OTHER INFORMATION: Exon 5</li></ul>	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:	
15	TCCTGTTGTT CTACAATGTA CACATGTAAC ACCACAAAGA GATAAGTCAG	50
	(2) INFORMATION FOR SEQ ID NO:2:	
20	<ul><li>(i) SEQUENCE CHARACTERISTICS:</li><li>(A) LENGTH: 182 base pairs</li><li>(B) TYPE: nucleic acid</li><li>(C) STRANDEDNESS: double</li><li>(D) TOPOLOGY: linear</li></ul>	
25	<pre>(ii) MOLECULE TYPE: Genomic DNA (ix) FEATURE:</pre>	
30	<ul><li>(A) NAME/KEY: exon</li><li>(B) LOCATION: 1182</li><li>(D) OTHER INFORMATION: Exon 15</li></ul>	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:	
35	ATTTAATTAC AAGTCTTCAG AATGCCAGAG ATATACAGGA TATGCGAATT AAGAAGAAAC AAAGGCAACG CGTCTTTCCA CAGCCAGGCA GTCTGTATCT TGCAAAAACA TCCACTCTGC CTCGAATCTC TCTGAAAGCA GCAGTAGGAG GCCAAGTTCC CTCTGCGTGT TCTCATAAAC AG	60 120 180 182
40	(2) INFORMATION FOR SEQ ID NO:3:	
45	<ul><li>(i) SEQUENCE CHARACTERISTICS:</li><li>(A) LENGTH: 188 base pairs</li><li>(B) TYPE: nucleic acid</li><li>(C) STRANDEDNESS: double</li><li>(D) TOPOLOGY: linear</li></ul>	
	(ii) MOLECULE TYPE: Genomic DNA (ix) FEATURE:	
50	<ul><li>(A) NAME/KEY: exon</li><li>(B) LOCATION: 1188</li><li>(D) OTHER INFORMATION: Exon 16</li></ul>	
55	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:	
	CTGTATACGT ATGGCGTTTC TAAACATTGC ATAAAAATTA ACAGCAAAAA TGCAGAGTCT TTTCAGTTTC ACACTGAAGA TTATTTTGGT AAGGAAAGTT TATGGACTGG AAAAGGAATA CAGTTGGCTG ATGGTGGATG GCTCATACCC TCCAATGATG GAAAGGCTGG AAAAGAAGAA TTTTATAG	60 120 180 188
60	(2) INFORMATION FOR SEC ID NO.4.	

5	<ul><li>(i) SEQUENCE CHARACTERISTICS:</li><li>(A) LENGTH: 10485 base pairs</li><li>(B) TYPE: nucleic acid</li><li>(C) STRANDEDNESS: single</li><li>(D) TOPOLOGY: linear</li></ul>	
10	<ul><li>(ii) MOLECULE TYPE: cDNA</li><li>(ix) FEATURE:</li><li>(A) NAME/KEY: Coding Sequence</li><li>(B) LOCATION: 22910482</li></ul>	
15	(D) OTHER INFORMATION: BRCA2 (OMI1)  (xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:	
20	GGTGGCGCGA GCTTCTGAAA CTAGGCGGCA GAGGCGGAGC CGCTGTGGCA CTGCTGCGCC TCTGCTGCGC CTCGGGTGTC TTTTGCGGCG GTGGGTCGCC GCCGGGAGAA GCGTGAGGGG ACAGATTTGT GACCGGCGCG GTTTTTGTCA GCTTACTCCG GCCAAAAAAG AACTGCACCT	60 120 180 237
25	GGA TCC AAA GAG AGG CCA ACA TTT TTT GAA ATT TTT AAG ACA CGC TGC Gly Ser Lys Glu Arg Pro Thr Phe Phe Glu Ile Phe Lys Thr Arg Cys 5 10 15	285
30	AAC AAA GCA GAT TTA GGA CCA ATA AGT CTT AAT TGG TTT GAA GAA CTT Asn Lys Ala Asp Leu Gly Pro Ile Ser Leu Asn Trp Phe Glu Glu Leu 20 25 30 35	333
35	TCT TCA GAA GCT CCA CCC TAT AAT TCT GAA CCT GCA GAA GAA TCT GAA Ser Ser Glu Ala Pro Pro Tyr Asn Ser Glu Pro Ala Glu Glu Ser Glu 40 45 50	381
	CAT AAA AAC AAC AAT TAC GAA CCA AAC CTA TTT AAA ACT CCA CAA AGG His Lys Asn Asn Asn Tyr Glu Pro Asn Leu Phe Lys Thr Pro Gln Arg 55 60 65	429
40	AAA CCA TCT TAT AAT CAG CTG GCT TCA ACT CCA ATA ATA TTC AAA GAG Lys Pro Ser Tyr Asn Gln Leu Ala Ser Thr Pro Ile Ile Phe Lys Glu 70 75 80	477
45	CAA GGG CTG ACT CTG CCG CTG TAC CAA TCT CCT GTA AAA GAA TTA GAT Gln Gly Leu Thr Leu Pro Leu Tyr Gln Ser Pro Val Lys Glu Leu Asp 85 90 95	525
50	AAA TTC AAA TTA GAC TTA GGA AGG AAT GTT CCC AAT AGT AGA CAT AAA Lys Phe Lys Leu Asp Leu Gly Arg Asn Val Pro Asn Ser Arg His Lys 100 115 110	573
55	AGT CTT CGC ACA GTG AAA ACT AAA ATG GAT CAA GCA GAT GAT GTT TCC Ser Leu Arg Thr Val Lys Thr Lys Met Asp Gln Ala Asp Asp Val Ser 120 125 130	621
23	TGT CCA CTT CTA AAT TCT TGT CTT AGT GAA AGT CCT GTT GTT CTA CAA Cys Pro Leu Leu Asn Ser Cys Leu Ser Glu Ser Pro Val Val Leu Gln 135 140 145	669
60	TGT ACA CAT GTA ACA CCA CAA AGA GAT AAG TCA GTG GTA TGT GGG AGT Cys Thr His Val Thr Pro Gln Arg Asp Lys Ser Val Val Cys Gly Ser 150 160	717

5			CAT His														765
10			GAA Glu														813
			TTA Leu														861
15			GAA Glu														909
20			AAA Lys 230														957
25			TTT Phe														1005
30			GCA Ala														1053
30			AGC Ser														1101
35	GAA Glu	GAT Asp	GAA Glu	GTA Val 295	TAT Tyr	GAA Glu	ACA Thr	GTT Val	GTA Val 300	GAT Asp	ACC Thr	TCT Ser	GAA Glu	GAA Glu 305	GAT Asp	AGT Ser	1149
40			TTA Leu 310														1197
45			AGC Ser														1245
50	GAA Glu 340	TGT Cys	GAA Glu	AAA Lys	TCT Ser	AAA Lys 345	AAC Asn	CAA Gln	GTG Val	AAA Lys	GAA Glu 350	AAA Lys	TAC Tyr	TCA Ser	TTT Phe	GTA Val 355	1293
	TCT Ser	GAA Glu	GTG Val	GAA Glu	CCA Pro 360	AAT Asn	GAT Asp	ACT Thr	GAT Asp	CCA Pro 365	TTA Leu	GAT Asp	TCA Ser	AAT Asn	GTA Val 370	GCA Ala	1341
55	CAT His	CAG Gln	AAG Lys	CCC Pro 375	TTT Phe	GAG Glu	AGT Ser	GGA Gly	AGT Ser 380	GAC Asp	AAA Lys	ATC Ile	TCC Ser	AAG Lys 385	GAA Glu	GTT Val	1389
60			TCT Ser 390														1437

5						GAG Glu											1485
J						GAA Glu 425											1533
10						ACT Thr											1581
15						AAG Lys											1629
20	Arg	Asp	Glu 470	Glu	Gln	CAT His	Leu	Glu 475	Ser	His	Thr	Asp	Cys 480	Ile	Leu	Ala	1677
25	Val	Lys 485	Gln	Ala	Ile	TCT Ser	Gly 490	Thr	Ser	Pro	Val	Ala 495	Ser	Ser	Phe	Gln	1725
	Gly 500	Ile	Lys	Lys	Ser	ATA Ile 505	Phe	Arg	Ile	Arg	Glu 510	Ser	Pro	Lys	Glu	Thr 515	1773
30	Phe	Asn	Ala	Ser	Phe 520	TCA Ser	Gly	His	Met	Thr 525	Asp	Pro	Asn	Phe	Lys 530	Lys	1821
35						GAA Glu											1869
40	Gln	Lys	Glu 550	Asp	Ser	TTA Leu	Cys	Pro 555	Asn	Leu	Ile	Asp	Asn 560	Gly	Ser	Trp	1917
45	Pro	Ala 565	Thr	Thr	Thr	CAG Gln	Asn 570	Ser	Val	Ala	Leu	Lys 575	Asn	Ala	Gly	Leu	1965
	Ile 580	Ser	Thr	Leu	Lys	AAG Lys 585	Lys	Thr	Asn	Lys	Phe 590	Ile	Tyr	Ala	Ile	His 595	2013
50	Asp	Glu	Thr	Ser	Tyr 600	AAA Lys	Gly	Lys	Lys	Ile 605	Pro	Lys	Asp	Gln	Lys 610	Ser	2061
55						TCA Ser											2109
60						AAT Asn											2157
	AAA	AGA	AGC	TGT	TCA	CAG	AAT	GAT	TCT	GAA	GAA	CCA	ACT	TTG	TCC	TTA	2205

	Lys	Arg 645	Ser	Cys	Ser	Gln	Asn 650	Asp	Ser	Glu	Glu	Pro 655	Thr	Leu	Ser	Leu	
5				TTT Phe													2253
10				AAT Asn													2301
15				AAG Lys 695													2349
20				TGC Cys													2397
20				TCA Ser													2445
25				CAT His													2493
30				CTT Leu													2541
35				TCC Ser 775													2589
40				TCA Ser													2637
40				GTT Val													2685
45				GCT Ala													2733
50				TAC Tyr													2781
55				ACA Thr 855													2829
60				TCA Ser													2877
				GAG Glu													2925

AAT CTT GCT TTA GGA AAT ACT AAG GAA CTT CAT GAA ACA GAC TTG ACT Asn Leu Ala Leu Gly Asn Thr Lys Glu Leu His Glu Thr Asp Leu Thr TGT GTA AAC GAA CCC ATT TTC AAG AAC TCT ACC ATG GTT TTA TAT GGA Cys Val Asn Glu Pro Ile Phe Lys Asn Ser Thr Met Val Leu Tyr Gly GAC ACA GGT GAT AAA CAA GCA ACC CAA GTG TCA ATT AAA AAA GAT TTG Asp Thr Gly Asp Lys Gln Ala Thr Gln Val Ser Ile Lys Lys Asp Leu GTT TAT GTT CTT GCA GAG GAG AAC AAA AAT AGT GTA AAG CAG CAT ATA Val Tyr Val Leu Ala Glu Glu Asn Lys Asn Ser Val Lys Gln His Ile AAA ATG ACT CTA GGT CAA GAT TTA AAA TCG GAC ATC TCC TTG AAT ATA Lys Met Thr Leu Gly Gln Asp Leu Lys Ser Asp Ile Ser Leu Asn Ile GAT AAA ATA CCA GAA AAA AAT AAT GAT TAC ATG AAC AAA TGG GCA GGA Asp Lys Ile Pro Glu Lys Asn Asn Asp Tyr Met Asn Lys Trp Ala Gly CTC TTA GGT CCA ATT TCA AAT CAC AGT TTT GGA GGT AGC TTC AGA ACA Leu Leu Gly Pro Ile Ser Asn His Ser Phe Gly Gly Ser Phe Arg Thr GCT TCA AAT AAG GAA ATC AAG CTC TCT GAA CAT AAC ATT AAG AAG AGC Ala Ser Asn Lys Glu Ile Lys Leu Ser Glu His Asn Ile Lys Lys Ser AAA ATG TTC TTC AAA GAT ATT GAA GAA CAA TAT CCT ACT AGT TTA GCT Lys Met Phe Phe Lys Asp Ile Glu Glu Gln Tyr Pro Thr Ser Leu Ala TGT GTT GAA ATT GTA AAT ACC TTG GCA TTA GAT AAT CAA AAG AAA CTG Cys Val Glu Ile Val Asn Thr Leu Ala Leu Asp Asn Gln Lys Lys Leu AGC AAG CCT CAG TCA ATT AAT ACT GTA TCT GCA CAT TTA CAG AGT AGT Ser Lys Pro Gln Ser Ile Asn Thr Val Ser Ala His Leu Gln Ser Ser GTA GTT GTT TCT GAT TGT AAA AAT AGT CAT ATA ACC CCT CAG ATG TTA Val Val Ser Asp Cys Lys Asn Ser His Ile Thr Pro Gln Met Leu TTT TCC AAG CAG GAT TTT AAT TCA AAC CAT AAT TTA ACA CCT AGC CAA Phe Ser Lys Gln Asp Phe Asn Ser Asn His Asn Leu Thr Pro Ser Gln AAG GCA GAA ATT ACA GAA CTT TCT ACT ATA TTA GAA GAA TCA GGA AGT Lys Ala Glu Ile Thr Glu Leu Ser Thr Ile Leu Glu Glu Ser Gly Ser CAG TTT GAA TTT ACT CAG TTT AGA AAA CCA AGC TAC ATA TTG CAG AAG 

Gln Phe Glu Phe Thr Gln Phe Arg Lys Pro Ser Tyr Ile Leu Gln Lys

5	AGT ACA TTT GAA GTG CCT GAA AAC CAG ATG ACT ATC TTA AAG ACC ACT Ser Thr Phe Glu Val Pro Glu Asn Gln Met Thr Ile Leu Lys Thr Thr 1140 1145 1150 1155	3693
10	TCT GAG GAA TGC AGA GAT GCT GAT CTT CAT GTC ATA ATG AAT GCC CCA Ser Glu Glu Cys Arg Asp Ala Asp Leu His Val Ile Met Asn Ala Pro 1160 1165 1170	3741
	TCG ATT GGT CAG GTA GAC AGC AGC CAA TTT GAA GGT ACA GTT GAA Ser Ile Gly Gln Val Asp Ser Ser Lys Gln Phe Glu Gly Thr Val Glu 1175 1180 1185	3789
15	ATT AAA CGG AAG TTT GCT GGC CTG TTG AAA AAT GAC TGT AAC AAA AGT Ile Lys Arg Lys Phe Ala Gly Leu Leu Lys Asn Asp Cys Asn Lys Ser 1190 1195 1200	3837
20	GCT TCT GGT TAT TTA ACA GAT GAA AAT GAA GTG GGG TTT AGG GGC TTT Ala Ser Gly Tyr Leu Thr Asp Glu Asn Glu Val Gly Phe Arg Gly Phe 1205 1210 1215	3885
25	TAT TCT GCT CAT GGC ACA AAA CTG AAT GTT TCT ACT GAA GCT CTG CAA Tyr Ser Ala His Gly Thr Lys Leu Asn Val Ser Thr Glu Ala Leu Gln 1220 1235 1230 1235	3933
30	AAA GCT GTG AAA CTG TTT AGT GAT ATT GAG AAT ATT AGT GAG GAA ACT Lys Ala Val Lys Leu Phe Ser Asp Ile Glu Asn Ile Ser Glu Glu Thr 1240 1245 1250	3981
30	TCT GCA GAG GTA CAT CCA ATA AGT TTA TCT TCA AGT AAA TGT CAT GAT Ser Ala Glu Val His Pro Ile Ser Leu Ser Ser Ser Lys Cys His Asp 1255 1260 1265	4029
35	TCT GTT GTT TCA ATG TTT AAG ATA GAA AAT CAT AAT GAT AAA ACT GTA Ser Val Val Ser Met Phe Lys Ile Glu Asn His Asn Asp Lys Thr Val 1270 1275 1280	4077
40	AGT GAA AAA AAT AAA TGC CAA CTG ATA TTA CAA AAT AAT ATT GAA Ser Glu Lys Asn Asn Lys Cys Gln Leu Ile Leu Gln Asn Asn Ile Glu 1285 1290 1295	4125
45	ATG ACT ACT GGC ACT TTT GTT GAA GAA ATT ACT GAA AAT TAC AAG AGA Met Thr Thr Gly Thr Phe Val Glu Glu Ile Thr Glu Asn Tyr Lys Arg 1300 1305 1310 1315	4173
50	AAT ACT GAA AAT GAA GAT AAC AAA TAT ACT GCT GCC AGT AGA AAT TCT Asn Thr Glu Asn Glu Asp Asn Lys Tyr Thr Ala Ala Ser Arg Asn Ser 1320 1325 1330	4221
30	CAT AAC TTA GAA TTT GAT GGC AGT GAT TCA AGT AAA AAT GAT ACT GTT His Asn Leu Glu Phe Asp Gly Ser Asp Ser Ser Lys Asn Asp Thr Val 1335 1340 1345	4269
55	TGT ATT CAT AAA GAT GAA ACG GAC TTG CTA TTT ACT GAT CAG CAC AAC Cys Ile His Lys Asp Glu Thr Asp Leu Leu Phe Thr Asp Gln His Asn 1350 1355 1360	4317
60	ATA TGT CTT AAA TTA TCT GGC CAG TTT ATG AAG GAG GGA AAC ACT CAG Ile Cys Leu Lys Leu Ser Gly Gln Phe Met Lys Glu Gly Asn Thr Gln 1365 1370 1375	4365

5	ATT AAA Ile Lys 1380			Leu					Phe					Lys		4413
_	CAA GAA Gln Glu		Cys					Ser					Leu			4461
10	ACT AAA Thr Lys	Thr					Lys					Ser				4509
15	TTT CAG Phe Gln					Lys					Ala					4557
20	AAT AAA Asn Lys 1445	Ile			Phe					Pro						4605
25	TTT TCC Phe Ser 1460			Ser					Asp					Lys		4653
	GAC ATT Asp Ile		Ser					Asp					Lys			4701
30	AAA GAA Lys Glu	Ser					Thr					Val				4749
35	GGA CAA Gly Gln					Glu					Pro					4797
40	TTT CAT Phe His 1525	Thr			Gly					Ile						4845
45	GAC AAA Asp Lys 1540	Val	Lys	Asn 1	Leu .545	Phe	Asp	Glu	Lys 1	Glu L550	Gln	Gly	Thr	Ser 1	Glu .555	4893
	ATC ACC	Ser	Phe 1	Ser .560	His	Gln	Trp	Ala 1	Lys .565	Thr	Leu	Lys	Tyr	Arg 1570	Glu	4941
50	GCC TGT Ala Cys	Lys					Ala					Glu				4989
55	GCC CCA Ala Pro					Met					Asn					5037
60	CTT GTT Leu Val 1605	Ser			Thr					Lys						5085
	TTA TGT	AGA	CAA	ACT	GAA	AAT	CTC	AAA	ACA	TCA	AAA	AGT	ATC	TTT	TTG	5133

	Leu 1620	Cys	Arg	Gln		Glu 1625	Asn	Leu	Lys		Ser 1630	Lys	Ser	Ile		Leu 1635	
5				GTA Val					Glu					Lys			5181
10			Cys	TAC Tyr 1655				Ser					Ile				5229
15		Leu		TTT Phe			Ser					Thr					5277
20	Thr			CTT Leu		Ala					Arg						5325
				GAA Glu	Arg					Asp					Tyr		5373
25				AAT Asn					Ile					Lys			5421
30			Glu	AAA Lys 1735				Tyr					Ser				5469
35		Tyr		TAC Tyr			Asp					Asp					5517
40	Ser			AAA Lys		Asp					Pro						5565
				AAA Lys	Asn					Lys					Val		5613
45				GCA Ala					Val					Cys			5661
50			Val	ACT Thr 1815				Pro					Asn				5709
55		Leu		ATA Ile			Ser					Val					5757
60	Phe			GCC Ala		Gly					Val						5805
30				AAA Lys													5853

CAA AAT GTA TCA AAA ATA CTT CCT CGT GTT GAT AAG AGA AAC CCA GAG

Gln Asn Val Ser Lys Ile Leu Pro Arg Val Asp Lys Arg Asn Pro Glu

	5				Asn					Lys					Glu	TTT Phe 2130	6621
				Asn					Glu					Glu		AAT Asn	6669
	10		Ile					Tyr					Gln			AAA Lys	6717
	15	Gln					Thr					Val				CAT His	6765
	20					Gln					Asn					ATT Ile	6813
Man is a sum had that the second	25				Thr					Pro					Ile	GAA Glu 2210	6861
America de de deservo	30			Thr					Ser					Glu		GAA Glu	6909
	30		Glu					Phe					Glu			GAT Asp	6957
	35	Lys					Ala					Phe				GAA Glu	7005
	40					Leu					Ile					GGA Gly	7053
	45				Leu	Val	Gly	Glu	Pro	Ser	Ile	Lys	Arg		Leu	TTA Leu 2290	7101
	50			Asp					Asn					Leu		GCT Ala	7149
	30		Ser					Thr					Arg			ATG Met	7197
	55	His					Pro					Pro				ACT Thr	7245
	60					Ile					Phe					CAA Gln	7293

_	TTT CTG TCT AAA TCT CAT TTG TAT GAA CAT CTG ACT TTG GAA AAA TCT  Phe Leu Ser Lys Ser His Leu Tyr Glu His Leu Thr Leu Glu Lys Ser  2360 2365 2370	341
5	TCA AGC AAT TTA GCA GTT TCA GGA CAT CCA TTT TAT CAA GTT TCT GCT  Ser Ser Asn Leu Ala Val Ser Gly His Pro Phe Tyr Gln Val Ser Ala  2375 2380 2385	389
10	ACA AGA AAT GAA AAA ATG AGA CAC TTG ATT ACT ACA GGC AGA CCA ACC  Thr Arg Asn Glu Lys Met Arg His Leu Ile Thr Thr Gly Arg Pro Thr  2390 2395 2400	437
15	AAA GTC TTT GTT CCA CCT TTT AAA ACT AAA TCA CAT TTT CAC AGA GTT  Lys Val Phe Val Pro Pro Phe Lys Thr Lys Ser His Phe His Arg Val  2405 2410 2415	485
20	GAA CAG TGT GTT AGG AAT ATT AAC TIG GAS GAN THE HOST GIN LYS GIN Glu Glu Cys Val Arg Asn Ile Asn Leu Glu Glu Asn Arg Gln Lys Gln 2420 2435	533
25	AAC ATT GAT GGA CAT GGC TCT GAT GAT AGT AAA AAT AAG ATT AAT GAC Asn Ile Asp Gly His Gly Ser Asp Asp Ser Lys Asn Lys Ile Asn Asp 2440 2445 2450	7581
23	AAT GAG ATT CAT CAG TIT AAC AAA AAC AAC ICC AAT CAN GON GON AS AS AS Glu Ile His Gln Phe As Lys As As Ser As Gln Ala Ala 2455 2460 2465	7629
30	Val Thr Phe Thr Lys Cys Glu Glu Glu Pro Leu Asp Leu Ile Thr Ser  2470 2475 2480	7677
35	Leu Gln Asn Ala Arg Asp Ile Gln Asp Met Arg Ile Lys Lys Gln 2485 2490 2495	7725
40	Arg Gln Arg Val Phe Pro Gln Pro Gly Ser Leu Tyr Leu Ala Lys Thr 2500 2505 2510 2515	7773
45	Ser Thr Leu Pro Arg Ile Ser Leu Lys Ala Ala Val Gly Gln Val 2520 2525 2530	7821
	Pro Ser Ala Cys Ser His Lys Gln Leu Tyr Thr Tyr Gly Val Ser Lys 2535 2540 2545	7869
50	His Cys Ile Lys Ile Asn Ser Lys Asn Ala Glu Ser Phe Gln Phe His 2550 2555 2560	7917
55	ACT GAA GAT TAT TTT GGT AAG GAA AGT TTA TGG ACT GGA AAA GGA ATA Thr Glu Asp Tyr Phe Gly Lys Glu Ser Leu Trp Thr Gly Lys Gly Ile 2565 2570 2575	7965
60	CAG TTG GCT GAT GGT GGA TGG CTC ATA CCC TCC AAT GAT GGA AAG GCT Gln Leu Ala Asp Gly Gly Trp Leu Ile Pro Ser Asn Asp Gly Lys Ala 2580 2585 2590 2595	8013
	GGA AAA GAA GAA TTT TAT AGG GCT CTG TGT GAC ACT CCA GGT GTG GAT	8061

	Gly Lys		Phe Tyr 2600	Arg Ala	Leu Cys 2605	Asp Thr		Val Asp 510	
5	CCA AAG Pro Lys	CTT ATT Leu Ile 2615	TCT AGA Ser Arg	Ile Trp	GTT TAT Val Tyr 620	AAT CAC Asn His	TAT AGA T Tyr Arg T 2625	IGG ATC Irp Ile	8109
10	Ile Trp	AAA CTG Lys Leu 2630	GCA GCT Ala Ala	ATG GAA Met Glu 2635	TGT GCC Cys Ala	TTT CCT Phe Pro	AAG GAA 1 Lys Glu 1 640	TTT GCT Phe Ala	8157
15	AAT AGA Asn Arg 2645	TGC CTA Cys Leu	Ser Pro	GAA AGG Glu Arg 2650	GTG CTT Val Leu	CTT CAA Leu Gln 2655	CTA AAA ' Leu Lys '	TAC AGA Tyr Arg	8205
20	TAT GAT Tyr Asp 2660	ACG GAA Thr Glu	ATT GAT Ile Asp 2665	AGA AGC Arg Ser	Arg Arg	TCG GCT Ser Ala 2670	ATA AAA A Ile Lys :	AAG ATA Lys Ile 2675	8253
20		Arg Asp				CTT GTT Leu Val	Leu Cys		8301
25	GAC ATA Asp Ile	ATT TCA Ile Ser 2695	TTG AGC Leu Ser	Ala Asn	ATA TCT Ile Ser 2700	GAA ACT Glu Thr	TCT AGC . Ser Ser . 2705	AAT AAA Asn Lys	8349
30	Thr Ser	AGT GCA Ser Ala 2710	GAT ACC Asp Thr	CAA AAA Gln Lys 2715	GTG GCC Val Ala	ATT ATT Ile Ile	GAA CTT . Glu Leu :720	ACA GAT Thr Asp	8397
35		Tyr Ala	Val Lys			CCT CCC Pro Pro 2735			8445
40					Gly Gln	AAG ATT Lys Ile 2750			8493
40		Leu Val				ACA CCT Thr Pro	Leu Glu		8541
45			Leu Lys	Ile Ser		AGT ACT Ser Thr			8589
50						CCT AGA Pro Arg			8637
55		Ser Ser				AAT GTT Asn Val 2815			8685
60				Tyr Pro		TGG ATG Trp Met 2830			8733
	TCT GGA Ser Gly	TTA TAC Leu Tyr	ATA TTT	CGC AAT Arg Asn	GAA AGA Glu Arc	GAG GAA Glu Glu	GAA AAG Glu Lys	GAA GCA Glu Ala	8781

2840 2845 2850

5	GCA AAA TAT GTG GAG GCC CAA CAA AAG AGA CTA GAA GCC TTA TTC ACT Ala Lys Tyr Val Glu Ala Gln Gln Lys Arg Leu Glu Ala Leu Phe Thr 2855 2860 2865	8829
10	AAA ATT CAG GAG GAA TTT GAA GAA CAT GAA GAA AAC ACA ACA AAA CCA Lys Ile Gln Glu Glu Phe Glu Glu His Glu Glu Asn Thr Thr Lys Pro 2870 2875 2880	8877
15	TAT TTA CCA TCA CGT GCA CTA ACA AGA CAG CAA GTT CGT GCT TTG CAA Tyr Leu Pro Ser Arg Ala Leu Thr Arg Gln Gln Val Arg Ala Leu Gln 2885 2890 2895	8925
13	GAT GGT GCA GAG CTT TAT GAA GCA GTG AAG AAT GCA GCA GAC CCA GCT Asp Gly Ala Glu Leu Tyr Glu Ala Val Lys Asn Ala Ala Asp Pro Ala 2900 2905 2910 2915	8973
20	TAC CTT GAG GGT TAT TTC AGT GAA GAG CAG TTA AGA GCC TTG AAT AAT Tyr Leu Glu Gly Tyr Phe Ser Glu Glu Gln Leu Arg Ala Leu Asn Asn 2920 2925 2930	9021
25	CAC AGG CAA ATG TTG AAT GAT AAG AAA CAA GCT CAG ATC CAG TTG GAA His Arg Gln Met Leu Asn Asp Lys Lys Gln Ala Gln Ile Gln Leu Glu 2935 2940 2945	9069
30	ATT AGG AAG GCC ATG GAA TCT GCT GAA CAA AAG GAA CAA GGT TTA TCA Ile Arg Lys Ala Met Glu Ser Ala Glu Gln Lys Glu Gln Gly Leu Ser 2950 2955 2960	9117
35	AGG GAT GTC ACA ACC GTG TGG AAG TTG CGT ATT GTA AGC TAT TCA AAA Arg Asp Val Thr Thr Val Trp Lys Leu Arg Ile Val Ser Tyr Ser Lys 2965 2970 2975	9165
35	AAA GAA AAA GAT TCA GTT ATA CTG AGT ATT TGG CGT CCA TCA TCA GAT Lys Glu Lys Asp Ser Val Ile Leu Ser Ile Trp Arg Pro Ser Ser Asp 2980 2985 2990 2995	9213
40	TTA TAT TCT CTG TTA ACA GAA GGA AAG AGA TAC AGA ATT TAT CAT CTT Leu Tyr Ser Leu Leu Thr Glu Gly Lys Arg Tyr Arg Ile Tyr His Leu 3000 3005 3010	9261
45	GCA ACT TCA AAA TCT AAA AGT AAA TCT GAA AGA GCT AAC ATA CAG TTA Ala Thr Ser Lys Ser Lys Ser Lys Ser Glu Arg Ala Asn Ile Gln Leu 3015 3020 3025	9309
50	GCA GCG ACA AAA AAA ACT CAG TAT CAA CAA CTA CCG GTT TCA GAT GAA Ala Ala Thr Lys Lys Thr Gln Tyr Gln Gln Leu Pro Val Ser Asp Glu 3030 3035 3040	9357
55	ATT TTA TTT CAG ATT TAC CAG CCA CGG GAG CCC CTT CAC TTC AGC AAA Ile Leu Phe Gln Ile Tyr Gln Pro Arg Glu Pro Leu His Phe Ser Lys 3045 3050 3055	9405
23	TTT TTA GAT CCA GAC TTT CAG CCA TCT TGT TCT GAG GTG GAC CTA ATA Phe Leu Asp Pro Asp Phe Gln Pro Ser Cys Ser Glu Val Asp Leu Ile 3060 3075	9453
60	GGA TTT GTC GTT TCT GTT GTG AAA AAA ACA GGA CTT GCC CCT TTC GTC Gly Phe Val Val Ser Val Val Lys Lys Thr Gly Leu Ala Pro Phe Val 3080 3085 3090	9501

5	TAT TTG TCA GAC GAA TGT TAC AAT TTA CTG GCA ATA AAG TTT TGG ATA Tyr Leu Ser Asp Glu Cys Tyr Asn Leu Leu Ala Ile Lys Phe Trp Ile 3095 3100 3105	9549
10	GAC CTT AAT GAG GAC ATT ATT AAG CCT CAT ATG TTA ATT GCT GCA AGC Asp Leu Asn Glu Asp Ile Ile Lys Pro His Met Leu Ile Ala Ala Ser 3110 3120	9597
10	AAC CTC CAG TGG CGA CCA GAA TCC AAA TCA GGC CTT CTT ACT TTA TTT Asn Leu Gln Trp Arg Pro Glu Ser Lys Ser Gly Leu Leu Thr Leu Phe 3125 3130 3135	9645
15	GCT GGA GAT TTT TCT GTG TTT TCT GCT AGT CCA AAA GAG GGC CAC TTT Ala Gly Asp Phe Ser Val Phe Ser Ala Ser Pro Lys Glu Gly His Phe 3140 3145 3150 3155	9693
20	CAA GAG ACA TTC AAC AAA ATG AAA AAT ACT GTT GAG AAT ATT GAC ATA Gln Glu Thr Phe Asn Lys Met Lys Asn Thr Val Glu Asn Ile Asp Ile 3160 3165 3170	9741
25	CTT TGC AAT GAA GCA GAA AAC AAG CTT ATG CAT ATA CTG CAT GCA AAT Leu Cys Asn Glu Ala Glu Asn Lys Leu Met His Ile Leu His Ala Asn 3175 3180 3185	9789
30	GAT CCC AAG TGG TCC ACC CCA ACT AAA GAC TGT ACT TCA GGG CCG TAC Asp Pro Lys Trp Ser Thr Pro Thr Lys Asp Cys Thr Ser Gly Pro Tyr 3190 3195 3200	9837
30	ACT GCT CAA ATC ATT CCT GGT ACA GGA AAC AAG CTT CTG ATG TCT TCT Thr Ala Gln Ile Ile Pro Gly Thr Gly Asn Lys Leu Leu Met Ser Ser 3205 3210 3215	9885
35	CCT AAT TGT GAG ATA TAT TAT CAA AGT CCT TTA TCA CTT TGT ATG GCC Pro Asn Cys Glu Ile Tyr Tyr Gln Ser Pro Leu Ser Leu Cys Met Ala 3220 3225 3230 3235	9933
40	AAA AGG AAG TCT GTT TCC ACA CCT GTC TCA GCC CAG ATG ACT TCA AAG Lys Arg Lys Ser Val Ser Thr Pro Val Ser Ala Gln Met Thr Ser Lys 3240 3245 3250	9981
45	TCT TGT AAA GGG GAG AAA GAG ATT GAT GAC CAA AAG AAC TGC AAA AAG Ser Cys Lys Gly Glu Lys Glu Ile Asp Asp Gln Lys Asn Cys Lys Lys 3255 3260 3265	10029
50	AGA AGA GCC TTG GAT TTC TTG AGT AGA CTG CCT TTA CCT CCA CCT GTT Arg Arg Ala Leu Asp Phe Leu Ser Arg Leu Pro Leu Pro Pro Pro Val 3270 3280	10077
30	AGT CCC ATT TGT ACA TTT GTT TCT CCG GCT GCA CAG AAG GCA TTT CAG Ser Pro Ile Cys Thr Phe Val Ser Pro Ala Ala Gln Lys Ala Phe Gln 3285 3290 3295	10125
55	CCA CCA AGG AGT TGT GGC ACC AAA TAC GAA ACA CCC ATA AAG AAA AAA Pro Pro Arg Ser Cys Gly Thr Lys Tyr Glu Thr Pro Ile Lys Lys 3300 3315	10173
60	GAA CTG AAT TCT CCT CAG ATG ACT CCA TTT AAA AAA TTC AAT GAA ATT Glu Leu Asn Ser Pro Gln Met Thr Pro Phe Lys Lys Phe Asn Glu Ile 3320 3325 3330	10221

5	Ser	Leu	Let	GAA Glu 3335	Ser	: AAT	TCA Ser	A ATA	A GCT Ala 3340	Asp	GAA Glu	GAA Glu	Leu	GCA Ala 3345	Lei	ATA l Ile	10269
3	AAT Asn	ACC Thr	CAA Glr 3350	ı Ala	CTI Leu	TTG Leu	TCI Ser	GGT Gly 3355	' Ser	ACA Thr	GGA Gly	GAA Glu	AAA Lys 3360	CAA Gln	TTI Phe	T ATA	10317
10	TCT Ser	GTC Val	Ser	GAA Glu	TCC Ser	ACT Thr	AGG Arg 3370	Thr	GCT Ala	CCC Pro	ACC Thr	AGT Ser 3375	TCA Ser	GAA Glu	GAT Asp	TAT Tyr	10365
15	CTC Leu 3380	Arg	CTG Leu	AAA Lys	Arg	CGT Arg 3385	Cys	' ACT Thr	' ACA Thr	TCT Ser	CTG Leu 3390	Ile	AAA Lys	GAA Glu	CAG Gln	GAG Glu 3395	10413
20	AGT Ser	TCC Ser	CAG Gln	Ala	AGT Ser 3400	Thr	GAA Glu	GAA Glu	Cys	GAG Glu 3405	Lys	AAT Asn	AAG Lys	Gln	GAC Asp 3410	ACA Thr	10461
25			Thr	AAA Lys 3415													10485
			(2	) IN	FORM	ATIO	N FO	R SE	Q ID	NO:	5 <b>:</b>						
30		(	i) S (A) (B) (C)	EQUE LENG TYPI STRA	NCE   GTH: E: al	CHAR 341: mino DNES:	ACTE 8 am aci	RIST ino d ingl	ICS: acid								
35				MOLE( RAGM)													
		(:	xi) :	SEQUI	ENCE	DES	CRIP	TION	: SE	Q ID	NO:	5:					
40	1			Gly	5					10					15	-	
				Asn 20					25					30	_		
45			35	Ser				40					45				
		50		His			55					60					
	65			Lys		70					75					80	
50				Gln	85					90					95	_	
				Lys 100					105					110			
55			115	Ser				120					125	Gln		_	
		130		Cys			135					140					
	Val 145	Leu	Gln	Cys	Thr	His 150	Val	Thr	Pro	Gln	Arg 155	Asp	Lys	Ser	Val	Val 160	
60	Cys	Gly	Ser	Leu	Phe 165	His	Thr	Pro	Lys	Phe 170		Lys	Gly	Arg	Gln 175	Thr	
	Pro	Lys	His	Ile	Ser	Glu	Ser	Leu	Gly		Glu	Val	qaA	Pro	Asp	Met.	

				180					185					190		
	Ser	Trp	Ser 195	Ser	Ser	Leu	Ala	Thr 200	Pro	Pro	Thr	Leu	Ser 205	Ser	Thr	Val
5	Leu	Ile 210	Val	Arg	Asn	Glu	Glu 215	Ala	Ser	Glu	Thr	Val 220	Phe	Pro	His	Asp
	Thr 225	Thr	Ala	Asn	Val	Lys 230	Ser	Tyr	Phe	Ser	Asn 235	His	Asp	Glu	Ser	Leu 240
10	Lys	Lys	Asn	Asp	Arg 245	Phe	Ile	Ala	Ser	Val 250	Thr	Asp	Ser	Glu	Asn 255	Thr
	Asn	Gln	Arg	Glu 260	Ala	Ala	Ser	His	Gly 265	Phe	Gly	Lys	Thr	Ser 270	Gly	Asn
	Ser	Phe	Lys 275	Val	Asn	Ser	Cys	Lys 280	Asp	His	Ile	Gly	Lys 285	Ser	Met	Pro
15	Asn	Val 290	Leu	Glu	Asp	Glu	Val 295	Tyr	Glu	Thr	Val	Val 300	Asp	Thr	Ser	Glu
	Glu 305	Asp	Ser	Phe	Ser	Leu 310	Cys	Phe	Ser	Lys	Cys 315	Arg	Thr	Lys	Asn	Leu 320
20	Gln	Lys	Val	Arg	Thr 325	Ser	Lys	Thr	Arg	Lys 330	Lys	Ile	Phe	His	Glu 335	Ala
	Asn	Ala	Asp	Glu 340	Cys	Glu	Lys	Ser	Lys 345	Asn	Gln	Val	Lys	Glu 350	Lys	Tyr
	Ser	Phe	Val 355	Ser	Glu	Val	Glu	Pro 360	Asn	Asp	Thr	Asp	Pro 365	Leu	Asp	Ser
25		370			Gln	_	375				_	380	_			
	Lys 385	Glu	Val	Val	Pro	Ser 390	Leu	Ala	Cys	Glu	Trp 395	Ser	Gln	Leu	Thr	Leu 400
30		_			Gly 405					410					415	
				420	Gln				425					430		
			435		Lys			440					445			
35		450			Pro	_	455		_			460				
	465		_		Asp -	470					475					480
40					Lys 485					490					495	
				500	Ile	_	_		505				_	510		
4 E	_		515					520					525			Asn
45		530			Thr		535					540				
	545	•			Lys	550	_			-	555				_	560
50	-		_		Ala 565 Ser					570					575	
				580	Glu				585					590		
55			595	_	Leu			600		_	_	_	605			_
J J		610			Leu		615	-				620				
	625					630					635		-			640
60				_	Arg 645 Ser					650	_				655	
	пец	Set	neu	660	SET	SET	± 11G	GIY	665	116	neu	ur A	nys	670	SET	νιά

	Asn	Glu	Thr 675	Cys	Ser	Asn	Asn	Thr 680	Val	Ile	Ser	Gln	Asp 685	Leu	Asp	Tyr
5	Lys	Glu 690	Ala	Lys	Cys	Asn	Lys 695	Glu	Lys	Leu	Gln	Leu 700	Phe	Ile	Thr	Pro
	Glu 705	Ala	Asp	Ser	Leu	Ser 710	Cys	Leu	Gln	Glu	Gly 715	Gln	Cys	Glu	Asn	Asp 720
	Pro	Lys	Ser	Lys	Lys 725	Val	Ser	Asp	Ile	Lys 730	Glu	Glu	Val	Leu	Ala 735	Ala
10	Ala	Cys	His	Pro 740	Val	Gln	His	Ser	Lys 745	Val	Glu	Tyr	Ser	Asp 750	Thr	Asp
		Gln	755		•			760	_	_			765			
15		Ile 770					775					780				
	785	Ser		_		790					795					800
		Asn	_		805	_				810					815	
20	-	Asn		820		_			825			_	_	830		
		Leu	835			-	-	840					845		_	_
25		Gln 850					855			_		860		_		
	865	Glu				870		-			875			_		880
30		Leu			885					890					895	
30		Arg Leu		900				_	905		_			910		
		Tyr	915					920					925			
35		930 Asp					935					940				
	945	_			_	950					955	_				960
4.0		His			965					970					975	
40		Asn		980	_				985			_	_	990		_
	_		995			_		1000	)				1005	5	_	
45		Arg 1010	)				1015	5				1020	)			
	Lys 1029	Lys 5	Ser	Lys	Met	Phe 1030		Lys	Asp	Ile	Glu 1035		Gln	Tyr	Pro	Thr 104
	Ser	Leu	Ala	Cys	Val 1045		Ile	Val	Asn	Thr 1050		Ala	Leu	Asp	Asn 1055	
50	Lys	Lys	Leu	Ser 1060		Pro	Gln	Ser	Ile 1065		Thr	Val	Ser	Ala 1070		Leu
		Ser	1075	5				1080	)				1085	5		
55		Met 1090	)			_	1099	5				1100	)			
	110					1110	)				1119	5				112
60		Gly			1125	5				1130	)				1135	5
00		Gln		1140	)				1145	5				1150	)	
	пув	Thr	1111	ser	GIU	GIU	Cys	Arg	Asp	Ala	Asp	ьeu	итѕ	vaı	тте	Met

		1155			1160					1165			
	Asn Ala	Pro Ser	Ile Gly		Val Z	Asp	Ser		Lys 1180		Phe	Glu	Gly
5	1185	Glu Ile	1190	С				1195					120
		Ser Ala	1205				1210					1215	5
10		Phe Tyr 1220	)			1225					1230	l	
		Gln Lys 1235			1240					1245	,		
	1250			1255					1260	)			
15	1265	Asp Ser	1270	0				1275	,				128
	_	Val Ser	1285				1290	)				1295	5
20		Glu Met 1300	)			1305					1310	)	
		Arg Asn 1315			1320					1325	5		
	1330			1335					1340	)			
25	1345	Val Cys	135	0				1355	5				136
		Asn Ile	1365				1370	)				1375	5
30		Gln Ile 138	0			1385	•				1390	)	
	_	Ala Gln 1395			1400					1405	5		
	141			1415	,				1420	0			
35	1425	Phe Phe	143	0				1435	5				144
		Phe Asn	1445				1450	)				145	5
40		Asn Phe 146	0			1465	5				1470	)	
	-	Met Asp 1475 Leu Lys			1480					1489	5		
4 =	149			1495	5				150	0			
45	1505	Gly Phe	151	0				1515	5				152
		Leu Asp	1525				1530	)				153	5
50		154 Glu Ile	0			1545	5				155	)	
		1555 Glu Ala			1560					156	5		
55	157			1575	5				158	0			
55	1585	Asn Leu	159	0				159	5				160
		Asn Leu	1605				1610	0				161	5
60	_	162 Leu Lys	0			1625	5				163	0	
	IIC FIIC	1635	· wr Dys	. , , ,	1640				J_ u	164			

	Lys	Ser 1650		Ala	Thr	Cys	Tyr 1655		Asn	Gln	Ser	Pro 1660		Ser	Val	Ile
	Glu	Asn	Ser	Ala	Leu	Ala	Phe	Tyr	Thr	Ser	Cys	Ser	Arg	Lys	Thr	Ser
5	1665					1670		-			1675		_	=		168
	Val	Ser	Gln	Thr	Ser	Leu	Leu	Glu	Ala	Lvs	Lvs	Trp	Leu	Arg	Glu	Glv
					1685					1690				3	1695	
	т10	Dho	7.00	C111			~1.u	720	Tlo			ב ד ת	Acn	Tyr		
	116	PHE	Asp	_		PIO	GIU				TIIL	Ата	Asp			Gry
1.0	_	_	_	1700		_	_		1705					1710		_
10	Asn	Tyr		_	Glu	Asn	Asn			Ser	Thr	ше		Glu	Asn	Asp
			1715					1720					1725			
	Lys	Asn	His	Leu	Ser	Glu	Lys	Gln	Asp	Thr	Tyr	Leu	Ser	Asn	Ser	Ser
		1730	)				1735	5				1740	)			
	Met	Ser	Asn	Ser	Tyr	Ser	Tyr	His	Ser	Asp	Glu	Val	Tyr	Asn	Asp	Ser
15	1745				_	1750	_			_	1755		_		_	176
			Leu	Ser	Lvs	Asn	Lvs	Leu	Asp	Ser	Glv	Ile	Glu	Pro	Val	Leu
	2	- 4 -			1765		_2 -			1770					1775	
	Lare	Λan	₹7a 1	Glu			Lare	Λen	Thr			Cor	Laze	Val		
	БУБ	ASII	vai	1780	_	OIII	БуЗ		1785		TITC	DCI	цур	1790		DCI
20	3	37-3	T			3	7.7				m l	17- 7	7			T1_
20	Asn	vaı	_	_	Ата	ASI	Ala			GIII	Thr	vaı		Glu	Asp	TIE
			1795			_	_	1800					1805			
	Cys	Val	Glu	Glu	Leu	Val			Ser	Ser	Pro	_	-	Asn	Lys	Asn
		1810	)				1815	5				1820	)			
	Ala	Ala	Ile	Lys	Leu	Ser	Ile	Ser	Asn	Ser	Asn	Asn	Phe	Glu	Val	Gly
25	1825	5				1830	)				1835	5				184
	Pro	Pro	Ala	Phe	Arq	Ile	Ala	Ser	Gly	Lys	Ile	Val	Cys	Val	Ser	His
					1845				•	1850			-		1855	
	Glu	Thr	Tle	Lve			Lve	Δsn	Tle			Asn	Ser	Phe		
	GIU	1111	110	1860	-	val	<b>-</b> y 5	_	1865			nop	001	1870		Lys
2.0	17_ T	T3 -	T			7	<b>a</b> 1				T	T1_	C			T
30	vaı	шe			Asn	ASI	GIU			ser	гàг	тте		Gln	Thr	гàг
	_		1875				_	1880					1885		_	
	Ile			GLY	Cys	Tyr			Leu	Asp	Asp			Asp	Ile	Leu
		1890					1895					1900				
	His	Asn	Ser	Leu	Asp	Asn	Asp	Glu	Cys	Ser	Thr	His	Ser	His	Lys	Val
35	1905	5				1910	)				1915	5				192
	Phe	Ala	Asp	Ile	Gln	Ser	Glu	Glu	Ile	Leu	Gln	His	Asn	Gln	Asn	Met
					1925	5				1930	)				1935	5
	Ser	Gly	Leu	Glu	Lys	Val	Ser	Lys	Ile	Ser	Pro	Cys	qaA	Val	Ser	Leu
		-		1940				4	1945			4	-	1950		
40	Glu	Thr	Ser			Cvs	Lazs	Cvs			Glv	LVS	Len	His		Ser
- 0	014	1111	1955	_	110	Cyb	<b>L</b> , 5	1960		110	0-7	_,5	1965		<i>-,</i> 5	001
	1707	C-~			7 0 0	Thr	Crra			Dho	Cor	Thr			C1	Tira
	val			Ala	ASII	TIIT	_	_	TIE	PILE	ser			Ser	GIY	пув
	_	1970			_	_	1975		_		_	1980		2		_,
			GIn	Val	Ser	_		Ser	Leu	Gin			Arg	Gln	Val	
45	1985					1990					1995					200
	Ser	Glu	Ile	Glu	Asp	Ser	Thr	-				~	Laze	Val	Leu	Phe
					I-	JUL	1111	ьуs	Gln	Val	Phe	Ser	шур			_
					2005		1111	ьуs	Gln	Val 2010		Ser	цур		2015	5
	Lys	Ser	Asn		2005	5				2010	)			Asn	2015	
	Lys	Ser	Asn		2009 His	5				2010 Thr	)				2015 Thr	
50	-			Glu 2020	2009 His	Ser	Asp	Gln	Leu 2025	2010 Thr	) Arg	Glu	Glu	Asn 2030	2015 Thr )	Ala
50	-		Thr	Glu 2020 Pro	2009 His	Ser	Asp	Gln Ile	Leu 2025 Ser	2010 Thr	) Arg	Glu	Glu Phe	Asn 2030 Ser	2015 Thr )	Ala
50	Ile	Arg	Thr 2035	Glu 2020 Pro	2005 His ) Glu	Ser His	Asp Leu	Gln Ile 2040	Leu 2025 Ser	2010 Thr Gln	Arg Lys	Glu Gly	Glu Phe 2045	Asn 2030 Ser	2019 Thr ) Tyr	Ala Asn
50	Ile	Arg Val	Thr 2035 Asn	Glu 2020 Pro	2005 His ) Glu	Ser His	Asp Leu Phe	Gln Ile 2040 Ser	Leu 2025 Ser	2010 Thr Gln	Arg Lys	Glu Gly Thr	Glu Phe 2045 Ala	Asn 2030 Ser	2019 Thr ) Tyr	Ala Asn
50	Ile Val	Arg Val 2050	Thr 2035 Asn	Glu 2020 Pro Ser	2009 His Glu Ser	Ser His Ala	Asp Leu Phe 2055	Gln Ile 2040 Ser	Leu 2025 Ser Gly	2010 Thr Gln Phe	Arg Lys Ser	Glu Gly Thr 2060	Glu Phe 2045 Ala	Asn 2030 Ser Ser	2019 Thr ) Tyr Gly	Ala Asn Lys
	Ile Val Gln	Arg Val 2050 Val	Thr 2035 Asn	Glu 2020 Pro Ser	2009 His Glu Ser	Ser His Ala Glu	Asp Leu Phe 2055 Ser	Gln Ile 2040 Ser	Leu 2025 Ser Gly	2010 Thr Gln Phe	Arg Lys Ser	Glu Gly Thr 2060 Val	Glu Phe 2045 Ala	Asn 2030 Ser	2019 Thr ) Tyr Gly	Ala Asn Lys Leu
50 55	Ile Val Gln 2069	Arg Val 2050 Val	Thr 2035 Asn Ser	Glu 2020 Pro Ser Ile	2009 His Glu Ser Leu	Ser His Ala Glu 2070	Asp Leu Phe 2055 Ser	Gln Ile 2040 Ser Ser	Leu 2025 Ser Gly Leu	2010 Thr Gln Phe	Arg Lys Ser Lys 2075	Glu Gly Thr 2060 Val	Glu Phe 2049 Ala ) Lys	Asn 2030 Ser Ser Gly	2019 Thr Tyr Gly Val	Ala Asn Lys Leu 208
	Ile Val Gln 2069	Arg Val 2050 Val	Thr 2035 Asn Ser	Glu 2020 Pro Ser Ile	2009 His Glu Ser Leu	Ser His Ala Glu 2070	Asp Leu Phe 2055 Ser	Gln Ile 2040 Ser Ser	Leu 2025 Ser Gly Leu	2010 Thr Gln Phe His	Arg Lys Ser Lys 2075 Ser	Glu Gly Thr 2060 Val	Glu Phe 2049 Ala ) Lys	Asn 2030 Ser Ser	2019 Thr Tyr Gly Val	Ala Asn Lys Leu 208
	Ile Val Gln 2069 Glu	Arg Val 2050 Val Glu	Thr 2035 Asn Ser Phe	Glu 2020 Pro Ser Ile Asp	Glu Ser Leu Leu 2089	Ser His Ala Glu 2070 Ile	Asp Leu Phe 2055 Ser Arg	Gln Ile 2040 Ser Ser Thr	Leu 2025 Ser Gly Leu Glu	2010 Thr Gln Phe His 2090	Lys Ser Lys Ser Ser Ser	Glu Gly Thr 2060 Val	Glu Phe 2049 Ala ) Lys His	Asn 2030 Ser Ser Gly	Thr Tyr Gly Val Ser	Ala Asn Lys Leu 208 Pro
	Ile Val Gln 2069 Glu	Arg Val 2050 Val Glu	Thr 2035 Asn Ser Phe	Glu 2020 Pro Ser Ile Asp	Glu Ser Leu Leu 2089	Ser His Ala Glu 2070 Ile	Asp Leu Phe 2055 Ser Arg	Gln Ile 2040 Ser Ser Thr	Leu 2025 Ser Gly Leu Glu	2010 Thr Gln Phe His 2090	Lys Ser Lys Ser Ser Ser	Glu Gly Thr 2060 Val	Glu Phe 2049 Ala ) Lys His	Asn 2030 Ser Ser Gly	Thr Tyr Gly Val Ser	Ala Asn Lys Leu 208 Pro
	Ile Val Gln 2069 Glu	Arg Val 2050 Val Glu	Thr 2035 Asn Ser Phe	Glu 2020 Pro Ser Ile Asp	2009 His Glu Ser Leu Leu 2089 Asn	Ser His Ala Glu 2070 Ile	Asp Leu Phe 2055 Ser Arg	Gln Ile 2040 Ser Ser Thr	Leu 2025 Ser Gly Leu Glu	2010 Thr Gln Phe His 2090 Leu	Lys Ser Lys Ser Ser Ser	Glu Gly Thr 2060 Val	Glu Phe 2049 Ala ) Lys His	Asn 2030 Ser Ser Gly	Thr Tyr Gly Val Ser 2099	Ala Asn Lys Leu 208 Pro
	Ile Val Gln 2069 Glu Thr	Arg Val 2050 Val Glu Ser	Thr 2035 Asn Ser Phe Arg	Glu 2020 Pro Ser Ile Asp Gln 2100	2005 His Glu Ser Leu Leu 2085 Asn	Ser His Ala Glu 2070 Ile Val	Asp Leu Phe 2055 Ser Arg	Gln Ile 2040 Ser Ser Thr	Leu 2025 Ser Gly Leu Glu Ile 2105	2010 Thr Gln Phe His 2090 Leu	Lys Ser Lys 2075 Ser	Glu Gly Thr 2060 Val Leu Arg	Glu Phe 2049 Ala Lys His	Asn 2030 Ser Ser Gly Tyr	Thr Tyr Gly Val Ser 2095	Ala Asn Lys Leu 208 Pro Arg
55	Ile Val Gln 2069 Glu Thr	Arg Val 2050 Val Glu Ser	Thr 2035 Asn Ser Phe Arg	Glu 2020 Pro Ser Ile Asp Gln 2100 His	2005 His Glu Ser Leu Leu 2085 Asn	Ser His Ala Glu 2070 Ile Val	Asp Leu Phe 2055 Ser Arg	Gln Ile 2040 Ser Ser Thr	Leu 2025 Ser Gly Leu Glu Ile 2105 Glu	2010 Thr Gln Phe His 2090 Leu	Lys Ser Lys 2075 Ser	Glu Gly Thr 2060 Val Leu Arg	Glu Phe 2049 Ala Lys His	Asn 2030 Ser Ser Gly Tyr Asp 2110 Cys	Thr Tyr Gly Val Ser 2095	Ala Asn Lys Leu 208 Pro Arg

		2130	)				2135					2140	)			
	Asn 2145	Asn	His	Ser	Ile	Lys 2150		Ser	Pro	Tyr	Leu 2155		Gln	Phe	Gln	Gln 216
5	-	Lys			2165	,		_		2170	)				2175	5
		His		2180	)	_			2185	5				2190	)	
10	Glu	Ile	Gly 2195	_	Thr	Glu	Thr	Phe 2200		Asp	Val	Pro	Val 2205	_	Thr	Asn
		Glu 2210	)	_			2215	,	_	_		2220	)			
	2225					2230	)				2235	,				224
15		Asp			2245	5				2250	1				2255	5
		Glu		2260	)				2265	5				2270	)	
20	_	Gly	2275	5				2280	)				2285	5		
		Leu 2290	)			_	2295	5				2300	)	_		
٥٢	2309			_		2310	)	_	_		2315	5	_			232
25		Met			2325	5				2330	)				2335	5
		Thr	-	2340	)				2345	5				2350	)	
30	_	Gln	2355	5			_	2360	)		_		2365	5		
		Lys 2370 Ser	)				2375	5				2380	)			
	2389	5			_	2390	)	_		_	2395	5				240
35	Arg	Pro	Thr	Lys	Val 2405		Val	Pro	Pro	Phe 2410		Thr	Lys	Ser	His 2415	
		Arg		2420	)	_		_	2425	5				2430	)	
40		Lys	2435	5				2440	)				2445	5		
	Ile	Asn 2450	_	Asn	Glu	Ile	His 2455		Phe	Asn	Lys	Asn 2460		Ser	Asn	Gln
	Ala 246		Ala	Val	Thr	Phe 2470		Lys	Cys	Glu	Glu 2475		Pro	Leu	Asp	Leu 248
45	Ile	Thr	Ser	Leu	Gln 2485		Ala	Arg	Asp	Ile 2490		Asp	Met	Arg	Ile 2495	
	Lys	Lys	Gln	Arg 2500	Gln		Val	Phe	Pro 2505	Gln		Gly	Ser	Leu 2510	_	Leu
50	Ala	Lys	Thr 2515		Thr	Leu	Pro	Arg 2520		Ser	Leu	Lys	Ala 2525		Val	Gly
	Gly	Gln 2530		Pro	Ser	Ala	Cys 2535		His	Lys	Gln	Leu 2540		Thr	Tyr	Gly
	254	5	_		_	2550	) _				2555	5				Phe 256
55					2565	5				2570	)				2575	
		Gly		2580	)				2585	5				2590	)	
60	_		259	5				2600	)	_			2605	5		Pro
	Gly	Val 261	_	Pro	гàг	ьeu	11e 2619		Arg	TIE	Trp	Val 2620	_	Asn	Hls	Tyr

	Arg Trp 2625	Ile I	Ile 7		Lys 2630		Ala	Ala		Glu 2635		Ala	Phe	Pro	Lys 264
_	Glu Phe	Ala A		Arg	Cys		Ser	Pro	Glu	Arg		Leu	Leu		Leu
5	Lys Tyr		Tyr A	2645 Asp		Glu	Ile	Asp 2665			Arg	Arg	Ser 2670		
	Lys Lys	Ile N	2660 Met (	Glu .	Arg	-	-	Thr		Ala	Lys	Thr 2685	Leu		Leu
10	Cys Val		Asp 3	Ile	Ile				Ala	Asn	Ile 2700	Ser		Thr	Ser
	Ser Asn 2705		Thr S		Ser 2710	Ala		Thr		Lys 2715	Val		Ile	Ile	Glu 272
15	Leu Thr	Asp (	_		Tyr		Val	Lys		Gln		Asp	Pro	Pro 2735	Leu
13	Leu Ala					Gly	_	Leu 2745	Thr		Gly	Gln	Lys 2750	Ile	
	Leu His			Glu	Leu			Ser		Asp	Ala	Cys 2765		Pro	Leu
20	Glu Ala 2770	Pro (	Glu s	Ser	Leu		Leu		Ile	Ser	Ala 2780		Ser	Thr	Arg
	Pro Ala 2785	Arg '	Trp ?	_	Thr 2790	_	Leu	Gly	Phe	Phe 2795		Asp	Pro	Arg	Pro 280
25	Phe Pro	Leu 1		Leu 2805		Ser	Leu	Phe	Ser 2810		Gly	Gly	Asn	Val 2815	
	Cys Val	_	Val : 2820	Ile	Ile	Gln	Arg	Ala 2825	-	Pro	Ile	Gln	Trp 2830		Glu
	Lys Thr	2835		_			2840	)				2845	5		
30	Lys Glu 285	C				2855	;				2860	)			
	Leu Phe 2865				2870	)				2875	5				288
35	Thr Lys			2885					2890	)				2895	5
	Ala Leu		2900	_				2905	5				2910	)	
	Asp Pro	2915	-			-	2920	)				2925	5	_	
40	Leu Asn 293	0		_		2935	5				2940	)			
	Gln Leu 2945			_	2950	)				2955	5				296
45	Gly Leu			2965					2970	)				2975	5
	Tyr Ser		2980					2985	5				2990	)	
Γ.Λ	Ser Ser	2995					3000	)				3005	5		
50	Tyr His	0				3015	5				3020	)			
	Ile Gln 3025				3030	)				3035	5				304
55	Ser Asp			3045	i				3050	)				305	5
	Phe Ser	_	3060					3065	5				3070	)	
60	Asp Leu	3075	_				3080	)				308	5		
00	Pro Phe 309 Phe Trp	0				3095	5				3100	)			
	rue 11p	тте	veh	⊔∈u	IIGA	GIU	App	116	116	пур	FIU	1112	riec	пец	116

	3105 3110 3115 312
	Ala Ala Ser Asn Leu Gln Trp Arg Pro Glu Ser Lys Ser Gly Leu Leu 3125 3130 3135
5	Thr Leu Phe Ala Gly Asp Phe Ser Val Phe Ser Ala Ser Pro Lys Glu 3140 3145 3150
	Gly His Phe Gln Glu Thr Phe Asn Lys Met Lys Asn Thr Val Glu Asn 3155 3160 3165
10	Ile Asp Ile Leu Cys Asn Glu Ala Glu Asn Lys Leu Met His Ile Leu 3170 3180
10	His Ala Asn Asp Pro Lys Trp Ser Thr Pro Thr Lys Asp Cys Thr Ser 3185 3190 3195 320
	Gly Pro Tyr Thr Ala Gln Ile Ile Pro Gly Thr Gly Asn Lys Leu Leu 3205 3210 3215
15	Met Ser Ser Pro Asn Cys Glu Ile Tyr Tyr Gln Ser Pro Leu Ser Leu 3220 3225 3230
	Cys Met Ala Lys Arg Lys Ser Val Ser Thr Pro Val Ser Ala Gln Met 3235 3240 3245
2.0	Thr Ser Lys Ser Cys Lys Gly Glu Lys Glu Ile Asp Asp Gln Lys Asn
20	3250 3255 3260  Cys Lys Lys Arg Arg Ala Leu Asp Phe Leu Ser Arg Leu Pro Leu Pro
	3265 3270 3275 328 Pro Pro Val Ser Pro Ile Cys Thr Phe Val Ser Pro Ala Ala Gln Lys
25	3285 3290 3295  Ala Phe Gln Pro Pro Arg Ser Cys Gly Thr Lys Tyr Glu Thr Pro Ile 3300 3305 3310
	Lys Lys Clu Leu Asn Ser Pro Gln Met Thr Pro Phe Lys Lys Phe
2.0	3315 3320 3325 Asn Glu Ile Ser Leu Leu Glu Ser Asn Ser Ile Ala Asp Glu Glu Leu
30	3330 3335 3340 Ala Leu Ile Asn Thr Gln Ala Leu Leu Ser Gly Ser Thr Gly Glu Lys
	3345 3350 3355 336  Gln Phe Ile Ser Val Ser Glu Ser Thr Arg Thr Ala Pro Thr Ser Ser
35	3365 3370 3375  Glu Asp Tyr Leu Arg Leu Lys Arg Arg Cys Thr Thr Ser Leu Ile Lys 3380 3385 3390
	Glu Gln Glu Ser Ser Gln Ala Ser Thr Glu Glu Cys Glu Lys Asn Lys
4.0	3395 3400 3405 Gln Asp Thr Ile Thr Thr Lys Lys Tyr Ile
40	3410 3415
	(2) INFORMATION FOR SEQ ID NO:6:
45	<ul><li>(i) SEQUENCE CHARACTERISTICS:</li><li>(A) LENGTH: 10485 base pairs</li><li>(B) TYPE: nucleic acid</li></ul>
	(C) STRANDEDNESS: double (D) TOPOLOGY: linear
50	(ii) MOLECULE TYPE: cDNA
~ ~	(ix) FEATURE:
	(A) NAME/KEY: Coding Sequence (B) LOCATION: 22910482
55	(D) OTHER INFORMATION: BRCA2 (OMI2)
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:
60	GGTGGCGCGA GCTTCTGAAA CTAGGCGGCA GAGGCGGAGC CGCTGTGGCA CTGCTGCGCC 6 TCTGCTGCGC CTCGGGTGTC TTTTGCGGCG GTGGGTCGCC GCCGGGAGAA GCGTGAGGGG 12
	ACAGATTTGT GACCGGCGC GTTTTTGTCA GCTTACTCCG GCCAAAAAAG AACTGCACCT 18

Met	Pro	Ile
1		

5		GAG Glu								:	285
10		GAT Asp								:	333
15		GCT Ala								•	381
2.0		AAC Asn 55						_			429
20		TAT Tyr									477
25		ACT Thr									525
30		TTA Leu									573
35		ACA Thr									621
4.0		CTA Leu 135									669
40		GTA Val									717
45		ACA Thr									765
50		AGT Ser							TCA Ser 195		813
55		GCT Ala									861
60		GAA Glu 215									909
60		AGC Ser							AAT Asn		957

GAT AGA TTT ATC GCT TCT GTG ACA GAC AGT GAA AAC ACA AAT CAA AGA Asp Arg Phe Ile Ala Ser Val Thr Asp Ser Glu Asn Thr Asn Gln Arg GAA GCT GCA AGT CAT GGA TTT GGA AAA ACA TCA GGG AAT TCA TTT AAA 

10				Ser												1033
15				TGC Cys												1101
13				GTA Val 295												1149
20				TGT Cys												1197
25				AAG Lys												1245
30				AAA Lys												1293
35				GAA Glu												1341
33				CCC Pro 375												1389
40				TTG Leu												1437
45				CAG Gln												1485
50	-	Gln	Asn	ATT Ile	Ser	Glu	Lys	Asp	Leu	Leu	Asp	Thr	Asn	Lys	Arg	1533
55				TTT Phe												1581
<i>J J</i>				TCA Ser 455												1629
60				GAG Glu												1677

5		AAG Lys 485								1725
10		ATC Ile								1773
		AAT Asn								1821
15		ACT Thr								1869
20		AAG Lys								1917
25		GCC Ala 565								1965
30	_	TCC Ser								2013
3 0		GAA Glu								2061
35		CTA Leu								2109
40		CTT Leu								2157
45		AGA Arg 645								2205
50		AGC Ser								2253
		TCT Ser								2301
55		TGT Cys								2349
60		CTG Leu								2397

5		AAA Lys 725															2445
J		GTA Val															2493
10		AAA Lys															2541
15		CCT Pro															2589
20		AAA Lys															2637
25		TCT Ser 805															2685
		GTA Val															2733
30		GAA Glu															2781
35		CAA Gln															2829
40		TCA Ser															2877
45		GAC Asp 885															2925
		CTT Leu															2973
50		GTA Val															3021
55		ACA Thr															3069
60		TAT Tyr															3117
	AAA	ATG	ACT	CTA	GGT	CAA	GAT	TTA	AAA	TCG	GAC	ATC	TCC	TTG	AAT	ATA	3165

	Lys Met Thr Leu Gly Gln Asp Leu Lys Ser Asp Ile Ser Leu Asn Ile 965 970 975	
5	GAT AAA ATA CCA GAA AAA AAT AAT GAT TAC ATG AAC AAA TGG GCA GGA Asp Lys Ile Pro Glu Lys Asn Asn Asp Tyr Met Asn Lys Trp Ala Gly 980 985 990 995	3213
10	CTC TTA GGT CCA ATT TCA AAT CAC AGT TTT GGA GGT AGC TTC AGA ACA Leu Leu Gly Pro Ile Ser Asn His Ser Phe Gly Gly Ser Phe Arg Thr 1000 1005 1010	3261
15	GCT TCA AAT AAG GAA ATC AAG CTC TCT GAA CAT AAC ATT AAG AAG AGC Ala Ser Asn Lys Glu Ile Lys Leu Ser Glu His Asn Ile Lys Lys Ser 1015 1020 1025	3309
20	AAA ATG TTC TTC AAA GAT ATT GAA GAA CAA TAT CCT ACT AGT TTA GCT Lys Met Phe Phe Lys Asp Ile Glu Glu Gln Tyr Pro Thr Ser Leu Ala 1030 1035 1040	3357
20	TGT GTT GAA ATT GTA AAT ACC TTG GCA TTA GAT AAT CAA AAG AAA CTG Cys Val Glu Ile Val Asn Thr Leu Ala Leu Asp Asn Gln Lys Lys Leu 1045 1050 1055	3405
25	AGC AAG CCT CAG TCA-ATT AAT ACT GTA TCT GCA CAT TTA CAG AGT AGT Ser Lys Pro Gln Ser Ile Asn Thr Val Ser Ala His Leu Gln Ser Ser 1060 1065 1070 1075	3453
30	GTA GTT GTT TCT GAT TGT AAA AAT AGT CAT ATA ACC CCT CAG ATG TTA Val Val Val Ser Asp Cys Lys Asn Ser His Ile Thr Pro Gln Met Leu 1080 1085 1090	3501
35	TTT TCC AAG CAG GAT TTT AAT TCA AAC CAT AAT TTA ACA CCT AGC CAA Phe Ser Lys Gln Asp Phe Asn Ser Asn His Asn Leu Thr Pro Ser Gln 1095 1100 1105	3549
40	AAG GCA GAA ATT ACA GAA CTT TCT ACT ATA TTA GAA GAA TCA GGA AGT Lys Ala Glu Ile Thr Glu Leu Ser Thr Ile Leu Glu Glu Ser Gly Ser 1110 1115 1120	3597
10	CAG TTT GAA TTT ACT CAG TTT AGA AAR CCA AGC TAC ATA TTG CAG AAG Gln Phe Glu Phe Thr Gln Phe Arg Xaa Pro Ser Tyr Ile Leu Gln Lys 1125 1130 1135	3645
45	AGT ACA TTT GAA GTG CCT GAA AAC CAG ATG ACT ATC TTA AAG ACC ACT Ser Thr Phe Glu Val Pro Glu Asn Gln Met Thr Ile Leu Lys Thr Thr 1140 1145 1150 1155	3693
50	TCT GAG GAA TGC AGA GAT GCT GAT CTT CAT GTC ATA ATG AAT GCC CCA Ser Glu Glu Cys Arg Asp Ala Asp Leu His Val Ile Met Asn Ala Pro 1160 1165 1170	3741
55	TCG ATT GGT CAG GTA GAC AGC AGC CAA TTT GAA GGT ACA GTT GAA Ser Ile Gly Gln Val Asp Ser Ser Lys Gln Phe Glu Gly Thr Val Glu 1175 1180 1185	3789
60	ATT AAA CGG AAG TTT GCT GGC CTG TTG AAA AAT GAC TGT AAC AAA AGT Ile Lys Arg Lys Phe Ala Gly Leu Leu Lys Asn Asp Cys Asn Lys Ser 1190 1195 1200	3837
	GCT TCT GGT TAT TTA ACA GAT GAA AAT GAA GTG GGG TTT AGG GGC TTT Ala Ser Gly Tyr Leu Thr Asp Glu Asn Glu Val Gly Phe Arg Gly Phe	3885

1205 1210 1215

5	TAT TCT GCT CAT GGC ACA AAA CTG AAT GTT TCT ACT GAA GCT CTG CAA Tyr Ser Ala His Gly Thr Lys Leu Asn Val Ser Thr Glu Ala Leu Gln 1220 1235	3933
10	AAA GCT GTG AAA CTG TTT AGT GAT ATT GAG AAT ATT AGT GAG GAA ACT Lys Ala Val Lys Leu Phe Ser Asp Ile Glu Asn Ile Ser Glu Glu Thr 1240 1245 1250	3981
15	TCT GCA GAG GTA CAT CCA ATA AGT TTA TCT TCA AGT AAA TGT CAT GAT Ser Ala Glu Val His Pro Ile Ser Leu Ser Ser Ser Lys Cys His Asp 1255 1260 1265	4029
	TCT GTC GTT TCA ATG TTT AAG ATA GAA AAT CAT AAT GAT AAA ACT GTA Ser Val Val Ser Met Phe Lys Ile Glu Asn His Asn Asp Lys Thr Val 1270 1280	4077
20	AGT GAA AAA AAT AAT AAA TGC CAA CTG ATA TTA CAA AAT AAT ATT GAA Ser Glu Lys Asn Asn Lys Cys Gln Leu Ile Leu Gln Asn Asn Ile Glu 1285 1290 1295	4125
25	ATG ACT ACT GGC ACT TTT GTT GAA GAA ATT ACT GAA AAT TAC AAG AGA Met Thr Thr Gly Thr Phe Val Glu Glu Ile Thr Glu Asn Tyr Lys Arg 1300 1305 1310 1315	4173
30	AAT ACT GAA AAT GAA GAT AAC AAA TAT ACT GCT GCC AGT AGA AAT TCT Asn Thr Glu Asn Glu Asp Asn Lys Tyr Thr Ala Ala Ser Arg Asn Ser 1320 1325 1330	4221
35	CAT AAC TTA GAA TTT GAT GGC AGT GAT TCA AGT AAA AAT GAT ACT GTT His Asn Leu Glu Phe Asp Gly Ser Asp Ser Ser Lys Asn Asp Thr Val 1335 1340 1345	4269
33	TGT ATT CAT AAA GAT GAA ACG GAC TTG CTA TTT ACT GAT CAG CAC AAC Cys Ile His Lys Asp Glu Thr Asp Leu Leu Phe Thr Asp Gln His Asn 1350 1355 1360	4317
40	ATA TGT CTT AAA TTA TCT GGC CAG TTT ATG AAG GAG GGA AAC ACT CAG Ile Cys Leu Lys Leu Ser Gly Gln Phe Met Lys Glu Gly Asn Thr Gln 1365 1370 1375	4365
45	ATT AAA GAA GAT TTG TCA GAT TTA ACT TTT TTG GAA GTT GCG AAA GCT Ile Lys Glu Asp Leu Ser Asp Leu Thr Phe Leu Glu Val Ala Lys Ala 1380 1385 1390 1395	4413
50	CAA GAA GCA TGT CAT GGT AAT ACT TCA AAT AAA GAA CAG TTA ACT GCT Gln Glu Ala Cys His Gly Asn Thr Ser Asn Lys Glu Gln Leu Thr Ala 1400 1405 1410	4461
55	ACT AAA ACG GAG CAA AAT ATA AAA GAT TTT GAG ACT TCT GAT ACA TTT Thr Lys Thr Glu Gln Asn Ile Lys Asp Phe Glu Thr Ser Asp Thr Phe 1415 1420 1425	4509
23	TTT CAG ACT GCA AGT GGG AAA AAT ATT AGT GTC GCC AAA GAG TCA TTT Phe Gln Thr Ala Ser Gly Lys Asn Ile Ser Val Ala Lys Glu Ser Phe 1430 1435 1440	4557
60	AAT AAA ATT GTA AAT TTC TTT GAT CAG AAA CCA GAA GAA TTG CAT AAC Asn Lys Ile Val Asn Phe Phe Asp Gln Lys Pro Glu Glu Leu His Asn 1445 1450 1455	4605

5					AAG AAC AAA ATG Lys Asn Lys Met 1475	4653
10		Leu Ser T		le Val Lys H	CAC AAA ATA CTG His Lys Ile Leu 1490	4701
					GTG ACC TTC CAG Wal Thr Phe Gln 1505	
15	Gly Gln		rg Asp Glu	ys Glu Pro T	ACT CTG TTG GGT Thr Leu Leu Gly 520	
20					AAG GAA TCT TTG Lys Glu Ser Leu	
25					GGT ACT AGT GAA Gly Thr Ser Glu 1555	
30		Ser Phe S		ys Thr Leu 1	AAG TAC AGA GAG Lys Tyr Arg Glu 1570	
30					GAG ATC ACA GCT Glu Ile Thr Ala 1585	
35	Ala Pro		ys Glu Met	er Leu Asn A	AAT GAT AAA AAC Asn Asp Lys Asn 600	
40					TTA AGT GAT AAT Leu Ser Asp Asn	
45					AGT ATC TTT TTG Ser Ile Phe Leu 1635	
50		Lys Val H		ys Glu Thr I	GCA AAA AGT CCT Ala Lys Ser Pro 1650	
30					ATT GAA AAT TCA Ile Glu Asn Ser 1665	
55	Ala Leu		Tyr Thr Ser	rg Lys Thr	TCT GTG AGT CAG Ser Val Ser Gln 680	
60		Leu Leu G			GGA ATA TTT GAT Gly Ile Phe Asp	

5		GCA GAT TAT GTA GGA AAT TAT TTG Ala Asp Tyr Val Gly Asn Tyr Leu 1710 1715	5373
~		ATA GCT GAA AAT GAC AAA AAT CAT Ile Ala Glu Asn Asp Lys Asn His 1725 1730	5421
10		TTA AGT AAC AGT AGC ATG TCT AAC Leu Ser Asn Ser Ser Met Ser Asn 1740 1745	5469
15		GTA TAT AAT GAT TCA GGA TAT CTC Val Tyr Asn Asp Ser Gly Tyr Leu 1760	5517
20		ATT GAG CCA GTA TTG AAG AAT GTT Ile Glu Pro Val Leu Lys Asn Val 1775	5565
25	Glu Asp Gln Lys Asn Thr Ser Phe 1780 1785	TCC AAA GTA ATA TCC AAT GTA AAA Ser Lys Val Ile Ser Asn Val Lys 1790 1795	5613
		GTA AAT GAA GAT ATT TGC GTT GAG Val Asn Glu Asp Ile Cys Val Glu 1805 1810	5661
30		TGC AAA AAT AAA AAT GCA GCC ATT Cys Lys Asn Lys Asn Ala Ala Ile 1820	5709
35		AAT TTT GAG GTA GGG CCA CCT GCA Asn Phe Glu Val Gly Pro Pro Ala 1840	5757
40		GTT TGT GTT TCA CAT GAA ACA ATT Val Cys Val Ser His Glu Thr Ile 1855	5805
45		A GAC AGT TTC AGT AAA GTA ATT AAG Asp Ser Phe Ser Lys Val Ile Lys 1870 1875	5853
	Glu Asn Asn Glu Asn Lys Ser Lys 1880	A ATT TGC CAA ACG AAA ATT ATG GCA B Ile Cys Gln Thr Lys Ile Met Ala 1885 1890	5901
50	Gly Cys Tyr Glu Ala Leu Asp Asp 1895	TCA GAG GAT ATT CTT CAT AAC TCT Ser Glu Asp Ile Leu His Asn Ser 1900 1905	5949
55	Leu Asp Asn Asp Glu Cys Ser Thr 1910 1915		5997
60		A CAT AAC CAA AAT ATG TCT GGA TTG A His Asn Gln Asn Met Ser Gly Leu 1935	6045
	GAG AAA GTT TCT AAA ATA TCA CC	TGT GAT GTT AGT TTG GAA ACT TCA	6093

	Glu 1940	Lys	Val	Ser		Ile 1945	Ser	Pro	Cys		Val 950	Ser	Leu	Glu	Thr 1	Ser .955	
5				Lys					Lys					Val	TCA Ser L970		6141
10			Thr					Ser					Lys		GTC Val		6189
15		Ser					Gln					Val			GAA Glu		6237
20	Glu					Gln					Val				AGT Ser		6285
20					Gln					Glu					CGT Arg		6333
25				Leu					Gly					Val	GTA Val 2050		6381
30			Ala					Ser					Lys		GTT Val		6429
35		Leu					His					Val			GAA Glu		6477
40	Asp					Glu					Tyr				TCT Ser		6525
40					Lys					Val					CCA Pro		6573
45				Asn					Lys					Glu	TTT Phe 2130		6621
50			Asn					Glu					Glu		AAT Asn		6669
55		Ile					Tyr					Gln			AAA Lys		6717
60	Gln					Thr					Val				CAT His		6765
															ATT Ile		6813

5		ATT Ile		Gly					Asp					Ile		7581
10		GAG Glu	Ile					Lys					Gln			7629
10		ACT Thr					Glu					Asp				7677
15	Leu	CAG Gln 2485				Asp					Arg					7725
20		CAA Gln			Phe					Ser					Lys	7773
25		ACT Thr		Pro					Lys					Gly		7821
30		TCT Ser	Ala					Gln					Gly			7869
30		TGC Cys					Ser					Ser				7917
35	Thr	GAA Glu 2565				Gly					Trp					7965
40		TTG Leu			Gly					Pro					Lys	8013
45		AAA Lys		Glu					Leu					Gly		8061
50		AAG Lys	Leu					Trp					Tyr			8109
30		TGG Trp					Met					Pro				8157
55	Asn	AGA Arg 2645				Pro	_		_		Leu	_				8205
60		GAT Asp			Ile					Arg					Lys	8253

5	ATG (			Asp					Lys					Cys			8301
5	GAC . Asp		Ile					Asn					Ser				8349
10	ACT . Thr	Ser					Gln					Ile					8397
15	GGG Gly 2					Lys					Pro						8445
20	TTA . Leu 2740				Arg					Gln					His		8493
25	GCA Ala			Val					Ala					Glu			8541
23	GAA Glu		Leu					Ser					Arg				8589
30	TGG Trp	Tyr					Phe					Arg					8637
35	CCC Pro 2					Phe					Asn						8685
40	GTA Val 2820				Arg					Gln					Thr		8733
45				Tyr				AAT Asn	Glu					Lys			8781
10	GCA Ala		Tyr					Gln					Ala				8829
50	AAA Lys	Ile					Glu					Asn					8877
55	Tyr					Ala		ACA Thr			Gln						8925
60					Leu			GCA Ala		Lys					Pro		8973
	TAC	CTT	GAG	GGT	TAT	TTC	AGT	GAA	GAG	CAG	TTA	AGA	GCC	TTG	AAT	AAT	9021

	Tyr Le	eu Gl	_	Tyr 2920	Phe	Ser	Glu		Gln 2925	Leu	Arg	Ala		Asn 2930	Asn	
5	CAC AG						Lys					Ile				9069
10	ATT AG		s Ala			Ser					Glu					9117
15	AGG GA Arg As 290	sp Va			Val					Ile						9165
20	AAA G Lys G 2980			Ser					Ile					Ser		9213
20	TTA TA		r Leu					Lys					Tyr			9261
25	GCA AG Ala Tl						Lys					Asn				9309
30	GCA GG Ala A		ır Lys			Gln					Pro					9357
35	ATT T Ile Lo 30	eu Ph			Tyr					Pro						9405
40	TTT T Phe Lo			Asp					Cys					Leu		9453
•	GGA T		al Val					Lys					Pro			9501
45	TAT T		CA GAC er Asp 3095				Asn					Lys				9549
50	GAC C' Asp L		n Glu			Ile					Leu					9597
55	AAC C' Asn Lo 31	eu Gl			Pro					Gly						9645
60	GCT G Ala G 3140			Ser					Ser					His		9693
	CAA G. Gln G		CA TTC ir Phe													9741

3160 3165 3170

5	Leu Cys Asn (	GAA GCA GAA AAC A Glu Ala Glu Asn 1 175			
10		TGG TCC ACC CCA F			
15		ATC ATT CCT GGT 7 Ile Ile Pro Gly ' 3210			
		GAG ATA TAT TAT ( Glu Ile Tyr Tyr ( 3225	Gln Ser Pro		
20		TCT GTT TCC ACA ( Ser Val Ser Thr : 3240		Ala Gln Met Thr	
25	Ser Cys Lys (	GGG GAG AAA GAG 2 Gly Glu Lys Glu 255			
30		TTG GAT TTC TTG . Leu Asp Phe Leu 3			
35		IGT ACA TTT GTT C Cys Thr Phe Val 3290			
33		AGT TGT GGC ACC . Ser Cys Gly Thr : 3305	Lys Tyr Glu		
40		TCT CCT CAG ATG . Ser Pro Gln Met ' 3320		Lys Lys Phe Asn	
45	Ser Leu Leu (	GAA AGT AAT TCA . Glu Ser Asn Ser 335			
50		GCT CTT TTG TCT ( Ala Leu Leu Ser ( 3			
55		GAA TCC ACT AGG . Glu Ser Thr Arg ' 3370			
33		AAA CGA CGT TGT . Lys Arg Arg Cys '	Thr Thr Ser		
60		GCC AGT ACG GAA Ala Ser Thr Glu 3400		Lys Asn Lys Gln	

## ATT ACA ACT AAA AAA TAT ATC TAA Ile Thr Thr Lys Lys Tyr Ile 3415

(2) INFORMATION FOR SEQ ID NO:7:

- 10 (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 3418 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

15

5

- (ii) MOLECULE TYPE: protein
  (v) FRAGMENT TYPE: internal
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

20			
<i>7</i> ()			
20			

Pro Gln Arg Lys Pro Ser Tyr Asn Gln Leu Ala Ser Thr Pro Ile Ile 30 65 70 75 80

Phe Lys Glu Gln Gly Leu Thr Leu Pro Leu Tyr Gln Ser Pro Val Lys

85 90 95
Glu Leu Asp Lys Phe Lys Leu Asp Leu Gly Arg Asn Val Pro Asn Ser
100 105 110

Arg His Lys Ser Leu Arg Thr Val Lys Thr Lys Met Asp Gln Ala Asp 115 120 125

Asp Val Ser Cys Pro Leu Leu Asn Ser Cys Leu Ser Glu Ser Pro Val 130 135 140

Val Leu Gln Cys Thr His Val Thr Pro Gln Arg Asp Lys Ser Val Val
40 145 150 155 160

Cys Gly Ser Leu Phe His Thr Pro Lys Phe Val Lys Gly Arg Gln Thr 165 170 175

Pro Lys His Ile Ser Glu Ser Leu Gly Ala Glu Val Asp Pro Asp Met
180 185 190

45 Ser Trp Ser Ser Ser Leu Ala Thr Pro Pro Thr Leu Ser Ser Thr Val \$195\$ 200 205

Leu Ile Val Arg Asn Glu Glu Ala Ser Glu Thr Val Phe Pro His Asp 210 215 220

Thr Thr Ala Asn Val Lys Ser Tyr Phe Ser Asn His Asp Glu Ser Leu 50 225 230 235 240

Lys Lys Asn Asp Arg Phe Ile Ala Ser Val Thr Asp Ser Glu Asn Thr 245 250 255

Asn Gln Arg Glu Ala Ala Ser His Gly Phe Gly Lys Thr Ser Gly Asn 260 265 270

Ser Phe Lys Val Asn Ser Cys Lys Asp His Ile Gly Lys Ser Met Pro 275 280 285

Asn Val Leu Glu Asp Glu Val Tyr Glu Thr Val Val Asp Thr Ser Glu 290 295 300

Glu Asp Ser Phe Ser Leu Cys Phe Ser Lys Cys Arg Thr Lys Asn Leu 305 310 315 320 Gln Lys Val Arg Thr Ser Lys Thr Arg Lys Lys Ile Phe His Glu Ala

325 330 335

	Asn	Ala	Asp	Glu 340	Cys	Glu	Lys	Ser	Lys 345	Asn	Gln	Val	Lys	Glu 350	Lys	Tyr
5	Ser	Phe	Val 355		Glu	Val	Glu	Pro 360		Asp	Thr	Asp	Pro 365		Asp	Ser
	Asn	Val 370		His	Gln	Lys	Pro 375		Glu	Ser	Gly	Ser 380		Lys	Ile	Ser
	Lys 385	Glu	Val	Val	Pro	Ser 390	Leu	Ala	Cys	Glu	Trp 395	Ser	Gln	Leu	Thr	Leu 400
10	Ser	Gly	Leu	Asn	Gly 405	Ala	Gln	Met	Glu	Lys 410	Ile	Pro	Leu	Leu	His 415	Ile
				420					425	_				430	Thr	
15		•	435	-	-	-		440					445		Pro	_
		450				_	455		_			460			Thr	
2.0	465		_		_	470					475				Asp	480
20					485					490					Ala 495 Ser	
				500		_	_		505				_	510	Pro	
25			515					520					525		His	
		530	_				535				_	540			Asp	
2.0	545					550					555					560
30	_		_		565					570					Lys 575	
				580					585					590	Ile	
35			595					600				_	605		Lys Asn	_
		610					615					620			Leu	
4.0	625					630					635		_			640
40					645		_			650					Pro 655	
				660					665					670	Ser	
45			675	-				680					685		Asp	-
		690					695					700			Thr	
50	705		_			710	_				715				Asn Ala	720
30				_	725			_		730					735	
				740					745			_		750	Thr	_
55			755		-			760	-	_			765		Val	
		770					775	_	_			780			Lys	
60	785			_	_	790		_			795	_	-		Met	800
00					805					810	_				815 Val	
	пув	TOIL	GIII	чор	vaı	CAR	мта	пеп	Mall	GIU	ASII	тАт	пур	Hall	vaı	GIU

				820					825					830		
		Leu	835			_		840					845			
5	Val	Gln 850	Phe	Asn	Gln	Asn	Thr 855	Asn	Leu	Arg	Val	Ile 860	Gln	Lys	Asn	Gln
	Glu 865	Glu	Thr	Thr	Ser	Ile 870	Ser	Lys	Ile	Thr	Val 875	Asn	Pro	Asp	Ser	Glu 880
10		Leu			885					890					895	
	Glu	Arg	Asn	Asn 900	Leu	Ala	Leu	Gly	Asn 905	Thr	Lys	Glu	Leu	His 910	Glu	Thr
	_	Leu	915	_				920			_		925			
15		Tyr 930	_			_	935					940				
	945	Asp			_	950					955	_				960
20		His			965			_		970					975	
		Asn		980	_				985			_		990		
0.5	_	Ala	995			_		1000	)				1005	5	_	
25		Arg 1010	)				1015	5		_		1020	)			
	102			_		1030	)				1035	5				104
30		Leu		_	1045	5				1050	)			_	1055	5
		Lys		1060	)				1065	5				1070	)	
2 5		Ser	1075	5				1080	)	_			1085	5		
35		Met 1090	)				1095	5				1100	)			
	110			-		1110	)				1115	5				112
40		Gly			1125	5				1130	)				1135	5
		Gln	_	1140		Pne	GIU	Val	PIO	(7   11	ASII	GIII	Met	TILL		Leu
4.5	Lys	Thr	Thr				_	_	1145	5				1150		
45	-		1155	5				1160	1145 Asp )	Ala	Asp	Leu	1165	Val	Ile	
		Ala 1170	1155 Pro )	Ser	Ile	Gly	Gln 1175	1160 Val	1145 Asp ) Asp	Ala Ser	Asp Ser	Leu Lys 1180	1165 Gln )	Val Dhe	Ile Glu	Gly
	Thr	1170 Val 5	1155 Pro ) Glu	Ser Ile	Ile Lys	Gly Arg 1190	Gln 1175 Lys	1160 Val S Phe	1145 Asp ) Asp Ala	Ala Ser Gly	Asp Ser Leu 1195	Leu Lys 1180 Leu	1165 Gln ) Lys	Val Phe Asn	Ile Glu Asp	Gly Cys 120
50	Thr 118 Asn	117( Val 5 Lys	1155 Pro ) Glu Ser	Ser Ile Ala	Ile Lys Ser 1205	Gly Arg 1190 Gly	Gln 1175 Lys ) Tyr	1160 Val Phe Leu	1145 Asp ) Asp Ala Thr	Ala Ser Gly Asp	Asp Ser Leu 1195 Glu	Leu Lys 1180 Leu Asn	1165 Gln ) Lys Glu	Val Phe Asn Val	Ile Glu Asp Gly 1215	Gly Cys 120 Phe
	Thr 118 Asn Arg	1170 Val 5 Lys Gly	Pro Glu Ser Phe	Ser Ile Ala Tyr 1220	Ile Lys Ser 1205 Ser	Gly Arg 1190 Gly Ala	Gln 1175 Lys ) Tyr His	1160 Val Fhe Leu	Asp Asp Ala Thr	Ala Ser Gly Asp 1210 Lys	Asp Ser Leu 1195 Glu Leu	Leu Lys 1180 Leu Asn	1165 Gln ) Lys Glu Val	Val Phe Asn Val Ser 1230	Ile Glu Asp Gly 1215 Thr	Gly Cys 120 Phe Glu
50	Thr 118 Asn Arg Ala	1170 Val 5 Lys Gly Leu	1155 Pro Glu Ser Phe Gln 1235	Ser Ile Ala Tyr 1220 Lys	Lys Ser 1205 Ser Ala	Gly Arg 1190 Gly Ala Val	Gln 1175 Lys ) Tyr His	1160 Val Phe Leu Gly Leu 1240	Asp Asp Ala Thr 1225 Phe	Ala Ser Gly Asp 1210 Lys Ser	Asp Ser Leu 1195 Glu Leu Asp	Leu Lys 1180 Leu Asn Asn	1165 Gln Lys Glu Val Glu 1245	Val Phe Asn Val Ser 1230 Asn	Ile Glu Asp Gly 1215 Thr	Cys 120 Phe Glu Ser
	Thr 118 Asn Arg Ala Glu	1170 Val 5 Lys Gly Leu Glu 1250	1155 Pro Glu Ser Phe Gln 1235 Thr	Ser Ile Ala Tyr 1220 Lys Ser	Lys Ser 1205 Ser Ala Ala	Gly Arg 1190 Gly Ala Val	Gln 1175 Lys ) Tyr His Lys Val 1255	1160 Val Phe Leu Gly Leu 1240 His	Asp Asp Asp Ala Thr Thr 1225 Phe Pro	Ala Ser Gly Asp 1210 Lys Ser Ile	Asp Ser Leu 1195 Glu Leu Asp	Leu Lys 1180 Leu Asn Asn Ile Leu 1260	1165 Gln Lys Glu Val Glu 1245 Ser	Val Phe Asn Val Ser 1230 Asn Ser	Glu Asp Gly 1215 Thr Ile Ser	Cys 120 Phe Glu Ser Lys
50	Thr 118 Asn Arg Ala Glu Cys 126	1170 Val 5 Lys Gly Leu Glu 1250 His	1155 Pro Glu Ser Phe Gln 1235 Thr	Ser Ile Ala Tyr 1220 Lys Ser Ser	Lys Ser 1205 Ser Ala Ala Val	Gly Arg 1190 Gly Ala Val Glu Val 1270	Gln 1175 Lys ) Tyr His Lys Val 1255 Ser	1160 Val Phe Leu Gly Leu 1240 His	Asp Asp Ala Thr Thr 1225 Phe Pro	Ala Ser Gly Asp 1210 Lys Ser Ile	Asp Ser Leu 1195 Glu Leu Asp Ser Ile 1275	Leu Lys 1180 Leu Asn Asn Ile Leu 1260 Glu	1165 Gln Lys Glu Val Glu 1245 Ser	Val Phe Asn Val Ser 1230 Asn Ser His	Glu Asp Gly 1215 Thr Ile Ser Asn	Cys 120 Phe Glu Ser Lys Asp 128
50	Thr 118 Asn Arg Ala Glu Cys 126 Lys	1170 Val 5 Lys Gly Leu Glu 1250 His	1155 Pro Olu Ser Phe Gln 1235 Thr Asp Val	Ser Ile Ala Tyr 1220 Lys Ser Ser	Lys Ser 1205 Ser Ala Ala Val Glu 1285	Gly Arg 1190 Gly Ala Val Glu Val 1270 Lys	Gln 1175 Lys Tyr His Val 1255 Ser	1160 Val Phe Leu Gly Leu 1240 His Met	Asp Asp Asp Ala Thr Thr 1225 Phe Pro Phe Lys	Ala Ser Gly Asp 1210 Lys Ser Ile Lys Cys 1290	Asp Ser Leu 1195 Glu Leu Asp Ser Ile 1275 Gln	Leu Lys 1180 Leu Asn Asn Leu 1260 Glu Leu	1165 Gln Lys Glu Val Glu 1245 Ser Asn	Val Phe Asn Val Ser 1230 Asn Ser His	Glu Asp Gly 1215 Thr Ile Ser Asn Gln 1295	Cys 120 Phe Glu Ser Lys Asp 128 Asn

	Tyr	Lys	Arg 1315		Thr	Glu	Asn	Glu 1320		Asn	Lys	Tyr	Thr 1325		Ala	Ser
5	Arg	Asn 1330		His	Asn	Leu	Glu 1335		Asp	Gly	Ser	Asp 1340		Ser	Lys	Asn
	Asp 1345	Thr	Val	Cys	Ile	His 1350		Asp	Glu	Thr	Asp 1355		Leu	Phe	Thr	Asp 136
	Gln	His	Asn	Ile	Cys 1365		Lys	Leu	Ser	Gly 1370		Phe	Met	Lys	Glu 1375	_
10	Asn	Thr	Gln	Ile 1380		Glu	Asp	Leu	Ser 1385		Leu	Thr	Phe	Leu 1390		Val
	Ala	Lys	Ala 1395		Glu	Ala	Cys	His 1400		Asn	Thr	Ser	Asn 1405		Glu	Gln
15	Leu	Thr 1410		Thr	Lys	Thr	Glu 1415		Asn	Ile	Lys	Asp 1420		Glu	Thr	Ser
	Asp 1425	Thr	Phe	Phe	Gln	Thr 1430		Ser	Gly	Lys	Asn 1435		Ser	Val	Ala	Lys 144
	Glu	Ser	Phe	Asn	Lys 1445		Val	Asn	Phe	Phe 1450		Gln	Lys	Pro	Glu 1455	
20		His		1460	)				1465	5				1470	)	
		Lys	1475	5				1480	)			_	1485	5	_	
25		Ile 1490	)				1495	5				1500	)			
	1509			_		1510	)	_	_		1515	5	_			152
		Leu			1525	5				1530	)				1535	5
30		Ser		1540	)				1545	5				1550	)	
		Ser	1555	5				1560	)		_		1565	5		_
35		Arg	)				1575	5				1580	)			
	1589					1590	)				1595	5				160
4.0	-	Lys			1605	5				1610	)			-	1615	õ
40		Asp Phe		1620	)				1625	5				1630	)	
			1635	5		_		1640	)				1645	5		
45	_	1650 Asn	)			_	1655	5				1660	) -			
	1665					1670	)				1675	5				168
50		Phe			1685	5				1690	)				1699	5
		Tyr		1700	)				1705	5				1710	)	
		Asn	1715	5				1720	)				1725	5		_
55	Met	1730 Ser		Ser	Tyr	Ser	1735 Tyr		Ser	Asp	Glu	1740 Val		Asn	Asp	Ser
•	1749 Gly	5 Tyr	Leu	Ser	Lys	1750 Asn		Leu	Asp	Ser	1759 Gly		Glu	Pro	Val	176 Leu
60	Lys	Asn	Val	Glu	1765 Asp		Lys	Asn	Thr	1770 Ser		Ser	Lys	Val	1775 Ile	
	Asn	Val	Lys	1780 Asp		Asn	Ala	Tyr	1785 Pro		Thr	Val	Asn	1790 Glu		Ile

			179	5				1800	١				1805	=		
	Cvs	Val		Glu	Leu	Val	Thr			Ser	Pro	Cvs			Lve	Aen
	-1-	181					1819		501	DCI	110	1820		Abii	цуз	ASII
5	Ala 1825		Ile	Lys	Leu	Ser 1830	Ile		Asn	Ser	Asn 1835	Asn		Glu	Val	Gly 184
			Ala	Phe	Arg 1845	Ile		Ser	Gly	Lys 1850	Ile		Cys	Val		His
10	Glu	Thr	Ile	Lys 1860	Lys		Lys	Asp	Ile 1869	Phe		Asp	Ser			
10	Val	Ile		Glu		Asn	Glu		Lys		Lys	Ile				Lys
	Ile			Gly	Cys	Tyr				Asp	Asp				Ile	Leu
15	uic	1890		Leu	7 an	Acn	1895		Cira	C 0 22	Th.	1900		TT-1	T	**- 7
13	1905		Ser	пец	Asp	1910		GIU	Cys	ser	1915		ser	HIS	гÀг	vai 192
	Phe	Ala	Asp	Ile	Gln 1925		Glu	Glu	Ile	Leu 1930		His	Asn	Gln	Asn 1935	Met
20	Ser	Gly	Leu	Glu 1940		Val	Ser	Lys	Ile 1945		Pro	Сув	Asp	Val 1950		Leu
	Glu	Thr	Ser 1955	Asp	Ile	Cys	Lys	Cys 1960		Ile	Gly	Lys	Leu 1965		Lys	Ser
		1970	)	Ala			1975	5				1980	)		-	-
25			Gln	Val	Ser			Ser	Leu	Gln			Arg	Gln	Val	Phe
	1985		Tlo	Glu	Nan	1990		Tira	C1 n	77-7	1995		T * * *	77- T	T	200
					2005	5				2010	)				2015	5
30				Glu 2020	)				2025	5				2030	1	
			2035					2040	)				2045	;		
	vaı	2050		Ser	ser	Ala	2055		GIY	Phe	Ser	Thr 2060		Ser	Gly	Lys
35	Gln 2065		Ser	Ile	Leu	Glu 2070		Ser	Leu	His	Lys 2075	Val		Gly	Val	Leu 208
	Glu	Glu	Phe	Asp	Leu 2085		Arg	Thr	Glu	His 2090		Leu	His	Tyr	Ser 2095	Pro
40	Thr	Ser	Arg	Gln 2100		Val	Ser	Lys	Ile 2105		Pro	Arg	Val	Asp 2110	Lys	
	Asn	Pro	Glu 2115	His	Cys	Val	Asn	Ser 2120		Met	Glu	Lys	Thr 2125	_	Ser	Lys
	Glu	Phe 2130		Leu	Ser	Asn	Asn 2135		Asn	Val	Glu	Gly 2140		Ser	Ser	Glu
45			His	Ser	Ile			Ser	Pro	Tyr			Gln	Phe	Gln	
	2145 Asp		Gln	Gln	Leu	2150 Val		Gly	Thr	Lvs	2155 Val		Leu	Val	Glu	216 Asn
					2165	;				2170	I				2175	;
50				Leu 2180	ı				2185					2190	_	
			2195					2200	)				2205			
		2210	)	Cys			2215	,				2220				
55	2225	5		Val		2230	1				2235					224
				Lys	2245					2250					2255	,
60				Glu 2260					2265					2270		
	Arg	Gly	Glu 2275	Pro	Leu	Ile	Leu	Val 2280		Glu	Pro		Ile 2285		Arg	Asn

	Leu	Leu 2290		Glu	Phe	Asp	Arg 229		Ile	Glu	Asn	Gln 2300		Lys	Ser	Leu
	Lvs	Ala	Ser	Lvs	Ser	Thr	Pro	Asp	Glv	Thr	Tle	Tivs	Asn	Ara	Ara	Len
5	2305					2310			1		2315		P		5	232
_			Hic	uic	Val			Glu	Dro	тіс			1/27	Dro	Dho	
	FIIC	Mec	111.5	1115	2325		пец	Giu	PIO	2330		Cys	vai	PIO	233	_
	Thr	Thr	Lys	Glu	Arg	Gln	Glu	Ile	Gln	Asn	Pro	Asn	Phe	Thr	Ala	Pro
				2340					2345					2350		
10	Gly	Gln	Glu 2355		Leu	Ser	Lys	Ser 2360		Leu	Tyr	Glu	His 2369	Leu		Leu
	Glu	Lve			Ser	Δen	T.=11			Sar	Gly	цiс			Фътъ	Cl n
	O I U	2370		DCI	DCI	ASII	2375		vai	Ser	Gry	2380		PITE	TYL	GII
	17.0.7			TT 10	7	7. ~ ~			7.7 - L	7	TT -			ml	m)	~ 7
1 F			Ата	THE	Arg			гуѕ	мес	Arg			шe	Inr	Thr	
15	2385		1	_		2390		_	_		2395					240
	Arg	Pro	Thr	гàа	Val 2409		Val	Pro	Pro	Phe 2410		Thr	Lys	Ser	His 2419	
	Hie	Ara	Val	Glu	Gln		Val	λκα	Λen			Len	Clu	C1.,		
	1112	тч	Vai	2420		Cys	vai				ASII	цец	Giu			Arg
20	~1 <b>~</b>	T	a1			7	<b>a</b> 1		2425					2430		_
20	GIII	ьуѕ			Ile	Asp	GTA			ser	Asp	Asp			Asn	гàг
		_	2435			_		244(					2445			
	lle			Asn	Glu	Ile			Phe	Asn	Lys			Ser	Asn	Gln
		2450					2455					2460				
	Ala	Ala	Ala	Val	Thr	Phe	Thr	Lys	Cys	Glu	Glu	Glu	Pro	Leu	Asp	Leu
25	2465	5				2470	)				2475	5				248
	Ile	Thr	Ser	Leu	Gln	Asn	Ala	Arg	Asp	Ile	Gln	Asp	Met	Arq	Ile	Lys
					2485			_	_	2490		-		_	2495	_
	Lvs	Lvs	Gln	Ara	Gln	Ara	Val	Phe	Pro			Glv	Ser	Leu		
	4			2500		5			2505			017	201	2510		
30	Δla	LVS	Thr		Thr	Len	Pro	Δνα			T. 211	Tage	בות			മിച
_ •		-1-	2515			Leu	110	2520		DCI	псα	цуз	2525		vai	Оту
	Glv	Gln			Ser	λla	Cvc			Laze	Gla	Len			Th rac	<u>ماء،</u>
	Cry	2530		FIO	Ser	AIG	2535		HIS	пуъ	GIII		_	1111	TAT	GIY
	1707			TT	Crra	T1.			7		T	2540		<b>~</b> 1	<b>a</b>	<b>5</b> 1
35			гуя	HIS	Cys			TTE	ASI	ser			Ala	GIU	Ser	
33	2545			1	~ 3	2550		_,		_	2555		_	_	_	256
	GIN	Pne	HIS	Thr	Glu		Tyr	Pne	GIY			Ser	Leu	Trp		
	_				2565		_		_	2570		_			2575	5
	Lys	GIY	Ile		Leu	Ala	Asp	Gly	Gly	Trp	Leu	Ile	Pro	Ser	Asn	Asp
	_			2580					2585					2590		
40	Gly	Lys			Lys	Glu	Glu			Arg	Ala	Leu	Cys	Asp	Thr	Pro
			2595					2600					2605			
	Gly	Val	Asp	Pro	Lys	Leu	Ile	Ser	Arg	Ile	Trp	Val	Tyr	Asn	His	Tyr
		2610	)				2615	5				2620	)			
	Arg	Trp	Ile	Ile	Trp	Lys	Leu	Ala	Ala	Met	Glu	Cys	Ala	Phe	Pro	Lys
45	2625					2630					2635					264
	Glu	Phe	Ala	Asn	Arg	Cys	Leu	Ser	Pro	Glu	Arq	Val	Leu	Leu	Gln	
					2645					2650					2655	
	Lvs	Tvr	Ara	Tvr	Asp		Glu	Tle	Asp			Ara	Δra	Ser		
	-1	- 1 -	5	2660					2665			9	9	2670		110
50	Lvs	Lvs	Tle		Glu	Ara	Aen	Aen			777	Larg	Thr			T 011
50	<i></i> , <i>-</i>	шуы	2675		GIU	Arg	дод	2680		Ala	Ата	цуъ			vai	ьеи
	Cvc	Val.			T10	Tlo	C 0 x			77-	7	T1 -	2685		m)	<b>a</b>
	СуБ			Asp	Ile	116			ser	Ala	ASII			GIU	Thr	ser
	0	2690		m1	•	_	2695		-1	~ 3	_	2700				
e e			ьуs	Thr	Ser			Asp	Thr	GIn			Ala	Ile	Ile	
55	2705		_			2710		_			2715					272
	Leu	Thr	Asp	GLY	Trp		Ala	Val	Lys	Ala	Gln	Leu	Asp	Pro	Pro	Leu
					2725					2730					2735	
	Leu	Ala	Val	Leu	Lys	Asn	Gly	Arg	Leu	Thr	Val	Gly	Gln	Lys	Ile	Ile
				2740					2745					2750		
60	Leu	His	Gly	Ala	Glu	Leu	Val	Gly	Ser	Pro	Asp	Ala	Cys	Thr	Pro	Leu
			2755					2760					2765			
	Glu	Ala	Pro	Glu	Ser	T.e.11	Met	T 11	Taze	TΤα	Sar	λlэ	) cn	Car	Thr	7~~

		2770	)				2775	5				2780	)			
	Pro 278	Ala	Arg	Trp	Tyr	Thr 2790		Leu	Gly	Phe	Phe 2799		Asp	Pro	Arg	Pro 280
5	Phe	Pro	Leu	Pro	Leu 2805		Ser	Leu	Phe	Ser 281		Gly	Gly	Asn	Val 281	_
		Val		2820	)				2825	5				2830	)	
10		Thr	2835	5				2840	)				2845	5		
		Glu 2850	)				2855	5				2860	)			
	2869					2870	)				2875	5				288
15		Lys			2885	5				2890	)				2895	5
		Leu		2900	)				2905	5				2910	)	
20		Pro	2915	5				2920	)				2925	5		
		Asn 2930	)				2935	5				2940	)			
2.5	2945					2950	)				2955	5				296
25		Leu			2965	5				2970	)		_		2975	5
		Ser		2980	)				2985	5				2990	)	
30		Ser	2995	5				3000	)				3005	5	_	
		His 3010 Gln	)				3015	5				3020	)			
	3025		пси	нта	AIG	3030		цуз	1111	GIII	3035		GIII	Leu	PIO	304
35		Asp			3045	5				3050	)				3055	5
		Ser		3060	)				3065	5				3070	)	
40		Leu	3075	5				3080	)				3085	i		
		Phe 3090	)				3095	5				3100	)			_
45	3105					3110	)				3115	i				312
45		Ala			3125	·				3130	)			_	3135	5
		Leu		3140	)				3145	i				3150	)	
50		His Asp	3155	5				3160	1				3165	·		
		3170 Ala	)				3175	;				3180	)			
55	3185					3190	)				3195	;				320
33		Ser			3205	;				3210	)				3215	5
		Met		3220	)				3225	;				3230	)	
60		Ser	3235	i				3240	)				3245	i		
	****	3250		<b>J</b> C1	Cys	ыyы	3255		ыys	JIU	116	3260		GIII	пув	HPII

	Cys Lys Lys Arg Arg Ala Leu Asp Phe Leu Ser Arg Leu Pro Leu Pro 3265 3270 3275 328	
	Pro Pro Val Ser Pro Ile Cys Thr Phe Val Ser Pro Ala Ala Gln Lys	
5	3285 3290 3295	
	Ala Phe Gln Pro Pro Arg Ser Cys Gly Thr Lys Tyr Glu Thr Pro Ile 3300 3305 3310	
	Lys Lys Lys Glu Leu Asn Ser Pro Gln Met Thr Pro Phe Lys Lys Phe 3315 3320 3325	
10	Asn Glu Ile Ser Leu Leu Glu Ser Asn Ser Ile Ala Asp Glu Glu Leu 3330 3335 3340	
	Ala Leu Ile Asn Thr Gln Ala Leu Leu Ser Gly Ser Thr Gly Glu Lys	
	3345 3350 3355 336 Gln Phe Ile Ser Val Ser Glu Ser Thr Arg Thr Ala Pro Thr Ser Ser	
15	3365 3370 3375	
	Glu Asp Tyr Leu Arg Leu Lys Arg Arg Cys Thr Thr Ser Leu Ile Lys 3380 3385 3390	
	Glu Gln Glu Ser Ser Gln Ala Ser Thr Glu Glu Cys Glu Lys Asn Lys	
20	3395 3400 3405 Gln Asp Thr Ile Thr Thr Lys Lys Tyr Ile	
	3410 3415	
	(2) INFORMATION FOR SEQ ID NO:8:	
25	(i) SEQUENCE CHARACTERISTICS:	
	(A) LENGTH: 10485 base pairs (B) TYPE: nucleic acid	
	(C) STRANDEDNESS: double	
30	(D) TOPOLOGY: linear	
30	(ii) MOLECULE TYPE: cDNA	
	(ix) FEATURE:	
	(A) NAME/KEY: Coding Sequence	
35	(B) LOCATION: 22910482	
	(D) OTHER INFORMATION: BRCA2 (OMI3)	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:	
40	GGTGGCGCGA GCTTCTGAAA CTAGGCGGCA GAGGCGGAGC CGCTGTGGCA CTGCTGCGCC	60
	TCTGCTGCGC CTCGGGTGTC TTTTGCGGCG GTGGGTCGCC GCCGGGAGAA GCGTGAGGGG ACAGATTTGT GACCGGCGCG GTTTTTGTCA GCTTACTCCG GCCAAAAAAG AACTGCACCT	120
		180 237
45	Met Pro Ile	
	1	
	GGA TCC AAA GAG AGG CCA ACA TTT TTT GAA ATT TTT AAG ACA CGC TGC	285
	Gly Ser Lys Glu Arg Pro Thr Phe Phe Glu Ile Phe Lys Thr Arg Cys 5 10 15	
50		
	AAC AAA GCA GAT TTA GGA CCA ATA AGT CTT AAT TGG TTT GAA GAA CTT Asn Lys Ala Asp Leu Gly Pro Ile Ser Leu Asn Trp Phe Glu Glu Leu	333
	20 25 30 35	
55	TCT TCA GAA GCT CCA CCC TAT AAT TCT GAA CCT GCA GAA GAA TCT GAA	201
	Ser Ser Glu Ala Pro Pro Tyr Asn Ser Glu Pro Ala Glu Glu Ser Glu	381
	40 45 50	
	CAT AAA AAC AAC AAT TAC GAA CCA AAC CTA TTT AAA ACT CCA CAA AGG	429
60	His Lys Asn Asn Asn Tyr Glu Pro Asn Leu Phe Lys Thr Pro Gln Arg	
	55 60 65	

5		CCA Pro															477
3		GGG Gly 85															525
10		TTC Phe															573
15		CTT Leu															621
20		CCA Pro															669
25	Cys	ACA Thr	His 150	Val	Thr	Pro	Gln	Arg 155	Asp	Lys	Ser	Val	Val 160	Cys	Gly	Ser	717
		TTT Phe 165															765
30		TCT Ser															813
35		TCT Ser															861
40	Arg	AAT Asn	Glu	Glu 215	Ala	Ser	Glu	Thr	Val 220	Phe	Pro	His	Asp	Thr 225	Thr	Ala	909
45	_	GTG Val															957
		AGA Arg 245															1005
50		GCT Ala															1053
55	Val	AAT Asn	Ser	Cys	Lys 280	Asp	His	Ile	Gly	Lys 285	Ser	Met	Pro	His	Val 290	Leu	1101
60		GAT Asp															1149
	TTT	TCA	TTA	TGT	TTT	TCT	AAA	TGT	AGA	ACA	AAA	TAA	CTA	CAA	AAA	GTA	1197

	Phe	Ser	Leu 310	Cys	Phe	Ser	Lys	Cys 315	Arg	Thr	Lys	Asn	Leu 320	Gln	Lys	Val	
5		ACT Thr 325															1245
10		TGT Cys															1293
15		GAA Glu															1341
20		CAG Gln															1389
20		CCG Pro															1437
25		GGA Gly 405															1485
30		CAA Gln															1533
35		AAA Lys															1581
40		CCA Pro			_					_	_			_			1629
40		GAT Asp															1677
45		AAG Lys 485															1725
50		ATC Ile															1773
55		AAT Asn															1821
60		ACT Thr															1869
00		AAG Lys															1917

550 555 560

5	CCA Pro	GCC Ala 565	ACC Thr	ACC Thr	ACA Thr	CAG Gln	AAT Asn 570	TCT Ser	GTA Val	GCT Ala	TTG Leu	AAG Lys 575	AAT Asn	GCA Ala	GGT Gly	TTA Leu	1965
10	ATA Ile 580	TCC Ser	ACT Thr	TTG Leu	AAA Lys	AAG Lys 585	AAA Lys	ACA Thr	AAT Asn	AAG Lys	TTT Phe 590	ATT Ile	TAT Tyr	GCT Ala	ATA Ile	CAT His 595	2013
15	GAT Asp	GAA Glu	ACA Thr	TCT Ser	TAT Tyr 600	AAA Lys	GGA Gly	AAA Lys	AAA Lys	ATA Ile 605	CCG Pro	AAA Lys	GAC Asp	CAA Gln	AAA Lys 610	TCA Ser	2061
13	GAA Glu	CTA Leu	ATT Ile	AAC Asn 615	TGT Cys	TCA Ser	GCC Ala	CAG Gln	TTT Phe 620	GAA Glu	GCA Ala	AAT Asn	GCT Ala	TTT Phe 625	GAA Glu	GCA Ala	2109
20	CCA Pro	CTT Leu	ACA Thr 630	TTT Phe	GCA Ala	AAT Asn	GCT Ala	GAT Asp 635	TCA Ser	GGT Gly	TTA Leu	TTG Leu	CAT His 640	TCT Ser	TCT Ser	GTG Val	2157
25	AAA Lys	AGA Arg 645	AGC Ser	TGT Cys	TCA Ser	CAG Gln	AAT Asn 650	GAT Asp	TCT Ser	GAA Glu	GAA Glu	CCA Pro 655	ACT Thr	TTG Leu	TCC Ser	TTA Leu	2205
30	ACT Thr 660	AGC Ser	TCT Ser	TTT Phe	GGG Gly	ACA Thr 665	ATT Ile	CTG Leu	AGG Arg	AAA Lys	TGT Cys 670	TCT Ser	AGA Arg	AAT Asn	GAA Glu	ACA Thr 675	2253
35	TGT Cys	TCT Ser	AAT Asn	AAT Asn	ACA Thr 680	GTA Val	ATC Ile	TCT Ser	CAG Gln	GAT Asp 685	CTT Leu	GAT Asp	TAT Tyr	AAA Lys	GAA Glu 690	GCA Ala	2301
33	AAA Lys	TGT Cys	AAT Asn	AAG Lys 695	GAA Glu	AAA Lys	CTA Leu	CAG Gln	TTA Leu 700	TTT Phe	ATT Ile	ACC Thr	CCA Pro	GAA Glu 705	GCT Ala	GAT Asp	2349
40	TCT Ser	CTG Leu	TCA Ser 710	TGC Cys	CTG Leu	CAG Gln	GAA Glu	GGA Gly 715	CAG Gln	TGT Cys	GAA Glu	AAT Asn	GAT Asp 720	Pro	AAA Lys	AGC Ser	2397
45	AAA Lys	AAA Lys 725	Val	TCA Ser	GAT Asp	ATA Ile	AAA Lys 730	GAA Glu	GAG Glu	GTC Val	TTG Leu	GCT Ala 735	Ala	GCA Ala	TGT Cys	CAC His	2445
50	CCA Pro 740	Val	CAA Gln	CAC His	TCA Ser	AAA Lys 745	Val	GAA Glu	TAC Tyr	AGT Ser	GAT Asp 750	Thr	GAC Asp	TTT Phe	CAA Gln	TCC Ser 755	2493
55	CAG Gln	AAA Lys	AGT Ser	CTT Leu	TTA Leu 760	Tyr	GAT Asp	CAT	GAA Glu	AAT Asn 765	Ala	: AGC	ACT Thr	CTT Leu	ATT Ile 770	TTA Leu	2541
33	ACT Thr	CCT Pro	ACT Thr	TCC Ser 775	Lys	GAT Asp	GTT Val	CTG	TCA Ser 780	Asn	CTA Lev	GTC Val	ATG Met	ATT : Ile 785	Ser	AGA Arg	2589
60	GGC Gly	: AAA Lys	GAA Glu 790	ı Ser	TAC Tyr	Lys	ATC Met	TCA Ser 795	Asp	AAC Lys	G CTC	C AAA	GGT Gly 800	/ Asr	CAAT n Asr	TAT Tyr	2637

5	GAA Glu	TCT Ser 805	GAT Asp	GTT Val	GAA Glu	TTA Leu	ACC Thr 810	AAA Lys	AAT Asn	ATT Ile	CCC Pro	ATG Met 815	GAA Glu	AAG Lys	AAT Asn	CAA Gln	2685
10	GAT Asp 820	GTA Val	TGT Cys	GCT Ala	TTA Leu	AAT Asn 825	GAA Glu	AAT Asn	TAT Tyr	AAA Lys	AAC Asn 830	GTT Val	GAG Glu	CTG Leu	TTG Leu	CCA Pro 835	2733
10	CCT Pro	GAA Glu	AAA Lys	TAC Tyr	ATG Met 840	AGA Arg	GTA Val	GCA Ala	TCA Ser	CCT Pro 845	TCA Ser	AGA Arg	AAG Lys	GTA Val	CAA Gln 850	TTC Phe	2781
15	AAC Asn	CAA Gln	AAC Asn	ACA Thr 855	AAT Asn	CTA Leu	AGA Arg	GTA Val	ATC Ile 860	CAA Gln	AAA Lys	AAT Asn	CAA Gln	GAA Glu 865	GAA Glu	ACT Thr	2829
20	ACT Thr	TCA Ser	ATT Ile 870	TCA Ser	AAA Lys	ATA Ile	ACT Thr	GTC Val 875	AAT Asn	CCA Pro	GAC Asp	TCT Ser	GAA Glu 880	GAA Glu	CTT Leu	TTC Phe	2877
25	TCA Ser	GAC Asp 885	AAT Asn	GAG Glu	AAT Asn	AAT Asn	TTT Phe 890	GTC Val	TTC Phe	CAA Gln	GTA Val	GCT Ala 895	AAT Asn	GAA Glu	AGG Arg	AAT Asn	2925
2.0	AAT Asn 900	CTT Leu	GCT Ala	TTA Leu	GGA Gly	AAT Asn 905	ACT Thr	AAG Lys	GAA Glu	CTT Leu	CAT His 910	GAA Glu	ACA Thr	GAC Asp	TTG Leu	ACT Thr 915	2973
30	TGT Cys	GTA Val	AAC Asn	GAA Glu	CCC Pro 920	ATT Ile	TTC Phe	AAG Lys	AAC Asn	TCT Ser 925	ACC Thr	ATG Met	GTT Val	TTA Leu	TAT Tyr 930	GGA Gly	3021
35	GAC Asp	ACA Thr	GGT Gly	GAT Asp 935	AAA Lys	CAA Gln	GCA Ala	ACC Thr	CAA Gln 940	GTG Val	TCA Ser	ATT	AAA Lys	AAA Lys 945	GAT Asp	TTG Leu	3069
40				Leu					Lys			GTA Val		Gln		ATA Ile	3117
45	AAA Lys	ATG Met 965	Thr	CTA Leu	GGT Gly	CAA Gln	GAT Asp 970	Leu	AAA Lys	TCG Ser	GAC Asp	ATC Ile 975	TCC Ser	TTG Leu	AAT Asn	ATA Ile	3165
50	GAT Asp 980	Lys	ATA Ile	. CCA Pro	GAA Glu	AAA Lys 985	Asn	AAT Asn	GAT Asp	TAC Tyr	ATG Met	Asp	AAA Lys	TGG Trp	GCA Ala	GGA Gly 995	3213
30	CTC Leu	TTA Leu	GGT Gly	CCA	ATT Ile	Ser	AAT Asn	CAC His	AGT Ser	TTT Phe	: Gly	GGT Gly	AGC Ser	TTC Phe	AGA Arg	ACA Thr	3261
55	GCT Ala	TCA Ser	AAT Asn	AAG Lys 1015	Glu	ATC	: AAG : Lys	CTC Lev	TCT Ser 1020	Glu	CAT His	AAC Asn	ATT	AAG Lys	Lys	AGC Ser	3309
60	AAA Lys	ATO Met	TTC Phe	Phe	: AAA : Lys	A GAT S Asp	TATI	GAA Glu 1035	ı Glu	CAA Glr	A TAT	CCT Pro	ACT Thr	Ser	TTA Leu	GCT Ala	3357

5	Cys	GTT Val 1045				Asn					Asp						3405
3	AGC Ser 1060	AAG Lys	CCT Pro	CAG Gln	Ser	ATT Ile .065	AAT Asn	ACT Thr	GTA Val	Ser	GCA Ala 1070	CAT His	TTA Leu	CAG Gln	Ser	AGT Ser .075	3453
10	GTA Val	GTT Val	GTT Val	Ser	GAT Asp 1080	TGT Cys	AAA Lys	AAT Asn	Ser	CAT His L085	ATA Ile	ACC Thr	CCT Pro	Gln	ATG Met L090	TTA Leu	3501
15	TTT Phe	TCC Ser	Lys	CAG Gln 1095	GAT Asp	TTT Phe	AAT Asn	Ser	AAC Asn 1100	CAT His	AAT Asn	TTA Leu	Thr	CCT Pro	AGC Ser	CAA Gln	3549
20		GCA Ala					Leu					Glu					3597
25	Gln	TTT Phe 1125				Gln					Ser						3645
	AGT Ser 1140	ACA Thr	TTT Phe	GAA Glu	Val	CCT Pro 1145	GAA Glu	AAC Asn	CAG Gln	Met	ACT Thr 1150	ATC Ile	TTA Leu	AAG Lys	Thr	ACT Thr 1155	3693
30		GAG Glu		Cys					Leu					Asn			3741
35	TCG Ser	ATT Ile	Gly	CAG Gln 1175	GTA Val	GAC Asp	AGC Ser	Ser	AAG Lys 1180	CAA Gln	TTT Phe	GAA Glu	Gly	ACA Thr 1185	GTT Val	GAA Glu	3789
40		AAA Lys					Gly					Asp					3837
45	Ala	TCT Ser 1205	Gly	Tyr	Leu	Thr	Asp 1210	Glu	Asn	Glu	Val	Gly 1215	Phe	Arg	Gly	Phe	3885
	Tyr 1220		Ala	His	Gly	Thr 1225	Lys	Leu	Asn	Val	Ser 1230	Thr	Glu	Ala	Leu	Gln 1235	3933
50	AAA Lys	A GCT s Ala	GTG Val	Lys	CTG Leu 1240	Phe	AGT Ser	GAT Asp	ATT Ile	GAG Glu 1245	Asn	ATT Ile	AGT Ser	GAG Glu	GAA Glu 1250	ACT Thr	3981
55		r GCA Ala			His			Ser		Ser					His		4029
60		r GTI c Val		Ser					Glu			Asn		Lys		GTA Val	4077
	AG:	r ga <i>f</i>	AAA A	LAA	TAA	' AAA	TGC	CAA	CTG	ATA	ATT A	CAA	raa .	' AA'	TTA	GAA	4125

	Ser Glu 1285	_	Asn	Asn		Cys .290	Gln	Leu	Ile		Gln .295	Asn	Asn	Ile	Glu	
5	ATG ACT Met Thr 1300			Thr					Ile					Lys		4173
10	AAT ACT Asn Thr		Asn					Tyr					Arg			4221
15	CAT AAC His Asn	Leu					Ser					Asn				4269
20	TGT ATT Cys Ile					Thr					Thr					4317
	ATA TGT Ile Cys 1365	Leu			Ser					Lys						4365
25	ATT AAA Ile Lys 1380			Leu					Phe					Lys		4413
30	CAA GAA Gln Glu		Cys					Ser					Leu			4461
35	ACT AAA Thr Lys	Thr					Lys					Ser				4509
40	TTT CAC					Lys					Ala					4557
	AAT AAA Asn Lys 1445	: Ile	Val	Asn	Phe	Phe	Asp	Gln	Lys	Pro	Glu	Glu				4605
45	TTT TCC Phe Ser 1460			Ser					Asp					Lys		4653
50	GAC ATT Asp Ile		Ser					Asp					Lys			4701
55	AAA GAA Lys Gli	ı Ser					Thr					Val				4749
60	GGA CAA Gly Gli	n Pro 1510	Glu	Arg	Asp	Glu	Lys 1515	Ile	Lys	Glu	Pro	Thr 1520	Leu	Leu	Gly	4797
	TTT CAT															4845

1525 1530 1535

5	GAC AAA GTG AAA AAC CTT TTT GAT GAA AAA GAG CAA GGT ACT AGT GAA Asp Lys Val Lys Asn Leu Phe Asp Glu Lys Glu Gln Gly Thr Ser Glu 1540 1545 1550 1555	4893
10	ATC ACC AGT TTT AGC CAT CAA TGG GCA AAG ACC CTA AAG TAC AGA GAG  Ile Thr Ser Phe Ser His Gln Trp Ala Lys Thr Leu Lys Tyr Arg Glu  1560 1565 1570	4941
1 5	GCC TGT AAA GAC CTT GAA TTA GCA TGT GAG ACC ATT GAG ATC ACA GCT Ala Cys Lys Asp Leu Glu Leu Ala Cys Glu Thr Ile Glu Ile Thr Ala 1575 1580 1585	4989
15	GCC CCA AAG TGT AAA GAA ATG CAG AAT TCT CTC AAT AAT GAT AAA AAC Ala Pro Lys Cys Lys Glu Met Gln Asn Ser Leu Asn Asn Asp Lys Asn 1590 1595 1600	5037
20	CTT GTT TCT ATT GAG ACT GTG GTG CCA CCT AAG CTC TTA AGT GAT AAT Leu Val Ser Ile Glu Thr Val Val Pro Pro Lys Leu Leu Ser Asp Asn 1605 1610 1615	5085
25	TTA TGT AGA CAA ACT GAA AAT CTC AAA ACA TCA AAA AGT ATC TTT TTG Leu Cys Arg Gln Thr Glu Asn Leu Lys Thr Ser Lys Ser Ile Phe Leu 1620 1625 1630 1635	5133
30	AAA GTT AAA GTA CAT GAA AAT GTA GAA AAA GAA ACA GCA AAA AGT CCT Lys Val Lys Val His Glu Asn Val Glu Lys Glu Thr Ala Lys Ser Pro 1640 1645 1650	5181
	GCA ACT TGT TAC ACA AAT CAG TCC CCT TAT TCA GTC ATT GAA AAT TCA Ala Thr Cys Tyr Thr Asn Gln Ser Pro Tyr Ser Val Ile Glu Asn Ser 1655 1660 1665	5229
35	GCC TTA GCT TTT TAC ACA AGT TGT AGT AGA AAA ACT TCT GTG AGT CAG Ala Leu Ala Phe Tyr Thr Ser Cys Ser Arg Lys Thr Ser Val Ser Gln 1670 1675 1680	5277
40	ACT TCA TTA CTT GAA GCA AAA AAA TGG CTT AGA GAA GGA ATA TTT GAT Thr Ser Leu Leu Glu Ala Lys Lys Trp Leu Arg Glu Gly Ile Phe Asp 1685 1690 1695	5325
45	GGT CAA CCA GAA AGA ATA AAT ACT GCA GAT TAT GTA GGA AAT TAT TTG Gly Gln Pro Glu Arg Ile Asn Thr Ala Asp Tyr Val Gly Asn Tyr Leu 1700 1705 1710 1715	5373
50	TAT GAA AAT AAT TCA AAC AGT ACT ATA GCT GAA AAT GAC AAA AAT CAT Tyr Glu Asn Asn Ser Asn Ser Thr Ile Ala Glu Asn Asp Lys Asn His 1720 1725 1730	5421
r.c	CTC TCC GAA AAA CAA GAT ACT TAT TTA AGT AAC AGT AGC ATG TCT AAC Leu Ser Glu Lys Gln Asp Thr Tyr Leu Ser Asn Ser Ser Met Ser Asn 1735 1740 1745	5469
55	AGC TAT TCC TAC CAT TCT GAT GAG GTA TAT AAT GAT TCA GGA TAT CTC Ser Tyr Ser Tyr His Ser Asp Glu Val Tyr Asn Asp Ser Gly Tyr Leu 1750 1755 1760	5517
60	TCA AAA AAT AAA CTT GAT TCT GGT ATT GAG CCA GTA TTG AAG AAT GTT Ser Lys Asn Lys Leu Asp Ser Gly Ile Glu Pro Val Leu Lys Asn Val 1765 1770 1775	5565

5	GAA GAT CAA AAA AAC ACT AGT TTT TCC AAA GTA ATA TCC AAT GTA AAA Glu Asp Gln Lys Asn Thr Ser Phe Ser Lys Val Ile Ser Asn Val Lys 1780 1785 1790 1795	5613
10	GAT GCA AAT GCA TAC CCA CAA ACT GTA AAT GAA GAT ATT TGC GTT GAG Asp Ala Asn Ala Tyr Pro Gln Thr Val Asn Glu Asp Ile Cys Val Glu 1800 1805 1810	5661
10	GAA CTT GTG ACT AGC TCT TCA CCC TGC AAA AAT AAA AAT GCA GCC ATT Glu Leu Val Thr Ser Ser Ser Pro Cys Lys Asn Lys Asn Ala Ala Ile 1815 1820 1825	5709
15	AAA TTG TCC ATA TCT AAT AGT AAT AAT TTT GAG GTA GGG CCA CCT GCA Lys Leu Ser Ile Ser Asn Ser Asn Asn Phe Glu Val Gly Pro Pro Ala 1830 1835 1840	5757
20	TTT AGG ATA GCC AGT GGT AAA ATC GTT TGT GTT TCA CAT GAA ACA ATT Phe Arg Ile Ala Ser Gly Lys Ile Val Cys Val Ser His Glu Thr Ile 1845 1850 1855	5805
25	AAA AAA GTG AAA GAC ATA TTT ACA GAC AGT TTC AGT AAA GTA ATT AAG Lys Lys Val Lys Asp Ile Phe Thr Asp Ser Phe Ser Lys Val Ile Lys 1860 1865 1870 1875	5853
	GAA AAC AAC GAG AAT AAA TCA AAA ATT TGC CAA ACG AAA ATT ATG GCA Glu Asn Asn Glu Asn Lys Ser Lys Ile Cys Gln Thr Lys Ile Met Ala 1880 1885 1890	5901
30	GGT TGT TAC GAG GCA TTG GAT GAT TCA GAG GAT ATT CTT CAT AAC TCT Gly Cys Tyr Glu Ala Leu Asp Asp Ser Glu Asp Ile Leu His Asn Ser 1895 1900 1905	5949
35	CTA GAT AAT GAT GAA TGT AGC ACG CAT TCA CAT AAG GTT TTT GCT GAC Leu Asp Asn Asp Glu Cys Ser Thr His Ser His Lys Val Phe Ala Asp 1910 1915 1920	5997
40	ATT CAG AGT GAA GAA ATT TTA CAA CAT AAC CAA AAT ATG TCT GGA TTG Ile Gln Ser Glu Glu Ile Leu Gln His Asn Gln Asn Met Ser Gly Leu 1925 1930 1935	6045
45	GAG AAA GTT TCT AAA ATA TCA CCT TGT GAT GTT AGT TTG GAA ACT TCA Glu Lys Val Ser Lys Ile Ser Pro Cys Asp Val Ser Leu Glu Thr Ser 1940 1945 1950 1955	6093
	GAT ATA TGT AAA TGT AGT ATA GGG AAG CTT CAT AAG TCA GTC TCA TCT Asp Ile Cys Lys Cys Ser Ile Gly Lys Leu His Lys Ser Val Ser Ser 1960 1965 1970	6141
50	GCA AAT ACT TGT GGG ATT TTT AGC ACA GCA AGT GGA AAA TCT GTC CAG Ala Asn Thr Cys Gly Ile Phe Ser Thr Ala Ser Gly Lys Ser Val Gln 1975 1980 1985	6189
55	GTA TCA GAT GCT TCA TTA CAA AAC GCA AGA CAA GTG TTT TCT GAA ATA Val Ser Asp Ala Ser Leu Gln Asn Ala Arg Gln Val Phe Ser Glu Ile 1990 1995 2000	6237
60	GAA GAT AGT ACC AAG CAA GTC TTT TCC AAA GTA TTG TTT AAA AGT AAC Glu Asp Ser Thr Lys Gln Val Phe Ser Lys Val Leu Phe Lys Ser Asn 2005 2010 2015	6285

5					Gln					Glu					CGT Arg 2		6333
5				Leu					Gly					Val	GTA Val		6381
10			Ala					Ser					Lys		GTT Val		6429
15		Leu					His					Val			GAA Glu		6477
20	Asp					Glu					Tyr				TCT Ser		6525
25					Lys					Val					CCA Pro		6573
23				Asn					Lys					Glu	TTT Phe 2130		6621
30			Asn					Glu					Glu		AAT Asn		6669
35		Ile					Tyr					Gln			AAA Lys		6717
40			Val			Thr					Val				CAT His		6765
45		Gly			Gln					Asn					ATT Ile		6813
40				Thr					Pro					Ile	GAA Glu 2210		6861
50			Thr		Ser			Ser					Glu		GAA Glu		6909
55		. Glu		Ala			Phe					Glu			GAT Asp		6957
60			Pro			Ala		His			Phe				GAA Glu		7005
	GAC	GAA	ATG	GTT	TTG	TCA	AAT	TCA	AGA	ATT	GGA	AAA	AGA	AGA	GGA	GAG	7053

	Glu Glu Met Val Leu Ser Asn Ser Arg Ile Gly Lys Arg Arg Gly Glu 2260 2265 2270 2275	
5	CCC CTT ATC TTA GTG GGA GAA CCC TCA ATC AAA AGA AAC TTA TTA AAT Pro Leu Ile Leu Val Gly Glu Pro Ser Ile Lys Arg Asn Leu Leu Asn 2280 2285 2290	7101
10	GAA TTT GAC AGG ATA ATA GAA AAT CAA GAA AAA TCC TTA AAG GCT TCA Glu Phe Asp Arg Ile Ile Glu Asn Gln Glu Lys Ser Leu Lys Ala Ser 2295 2300 2305	7149
15	AAA AGC ACT CCA GAT GGC ACA ATA AAA GAT CGA AGA TTG TTT ATG CAT Lys Ser Thr Pro Asp Gly Thr Ile Lys Asp Arg Arg Leu Phe Met His 2310 2315 2320	7197
20	CAT GTT TCT TTA GAG CCG ATT ACC TGT GTA CCC TTT CGC ACA ACT AAG His Val Ser Leu Glu Pro Ile Thr Cys Val Pro Phe Arg Thr Thr Lys 2325 2330 2335	7245
20	GAA CGT CAA GAG ATA CAG AAT CCA AAT TTT ACC GCA CCT GGT CAA GAA Glu Arg Gln Glu Ile Gln Asn Pro Asn Phe Thr Ala Pro Gly Gln Glu 2340 2345 2350 2355	7293
25	TTT CTG TCT AAA TCT CAT TTG TAT GAA CAT CTG ACT TTG GAA AAA TCT Phe Leu Ser Lys Ser His Leu Tyr Glu His Leu Thr Leu Glu Lys Ser 2360 2365 2370	7341
30	TCA AGC AAT TTA GCA GTT TCA GGA CAT CCA TTT TAT CAA GTT TCT GCT Ser Ser Asn Leu Ala Val Ser Gly His Pro Phe Tyr Gln Val Ser Ala 2375 2380 2385	7389
35	ACA AGA AAT GAA AAA ATG AGA CAC TTG ATT ACT ACA GGC AGA CCA ACC Thr Arg Asn Glu Lys Met Arg His Leu Ile Thr Thr Gly Arg Pro Thr 2390 2395 2400	7437
40	AAA GTC TTT GTT CCA CCT TTT AAA ACT AAA TCG CAT TTT CAC AGA GTT Lys Val Phe Val Pro Pro Phe Lys Thr Lys Ser His Phe His Arg Val 2405 2410 2415	7485
10	GAA CAG TGT GTT AGG AAT ATT AAC TTG GAG GAA AAC AGA CAA AAG CAA Glu Gln Cys Val Arg Asn Ile Asn Leu Glu Glu Asn Arg Gln Lys Gln 2420 2425 2430 2435	7533
45	AAC ATT GAT GGA CAT GGC TCT GAT GAT AGT AAA AAT AAG ATT AAT GAC Asn Ile Asp Gly His Gly Ser Asp Asp Ser Lys Asn Lys Ile Asn Asp 2440 2445 2450	7581
50	AAT GAG ATT CAT CAG TTT AAC AAA AAC AAC TCC AAT CAA GCA GCT Asn Glu Ile His Gln Phe Asn Lys Asn Asn Ser Asn Gln Ala Ala 2455 2460 2465	7629
55	GTA ACT TTC ACA AAG TGT GAA GAA GAA CCT TTA GAT TTA ATT ACA AGT Val Thr Phe Thr Lys Cys Glu Glu Pro Leu Asp Leu Ile Thr Ser 2470 2475 2480	7677
60	CTT CAG AAT GCC AGA GAT ATA CAG GAT ATG CGA ATT AAG AAG AAA CAA Leu Gln Asn Ala Arg Asp Ile Gln Asp Met Arg Ile Lys Lys Gln 2485 2490 2495	7725
- <b>-</b>	AGG CAA CGC GTC TTT CCA CAG CCA GGC AGT CTG TAT CTT GCA AAA ACA Arg Gln Arg Val Phe Pro Gln Pro Gly Ser Leu Tyr Leu Ala Lys Thr	7773

The state of the s

TTA AAG AAT GGC AGA CTG ACA GTT GGT CAG AAG ATT ATT CTT CAT GGA

Leu Lys Asn Gly Arg Leu Thr Val Gly Gln Lys Ile Ile Leu His Gly

5		Leu Val					CTT GAA GC Leu Glu Al 277	a Pro
10				Ile Ser			CGG CCT GC Arg Pro Al 2785	
10	Trp Tyr					Pro Arg	CCT TTT CC Pro Phe Pr 800	
15		Ser Ser	Leu Phe				GGT TGT GT Gly Cys Va	
20					Ile Gln		GAG AAG AC Glu Lys Th	
25		Leu Tyr					GAA AAG GA Glu Lys Gl 285	u Ala
2.0				Gln Gln			GCC TTA TT Ala Leu Ph 2865	
30						Glu Asn	ACA ACA AA Thr Thr Ly 880	
35		Pro Ser	Arg Ala				CGT GCT TT Arg Ala Le	
40					Val Lys		GCA GAC CC Ala Asp Pr	
45		ı Glu Gly					GCC TTG AA Ala Leu As 293	n Asn
50			Leu Asn	Asp Lys			ATC CAG TI Ile Gln Le 2945	
30						Lys Glu	CAA GGT TI Gln Gly Le 960	
55		Val Thr	Thr Val				AGC TAT TO Ser Tyr Se	
60				Ile Leu	Ser Ile		CCA TCA TO	

5	TTA TAT TCT CTG TTA ACA GAA GGA AAG AGA TAC AGA ATT TAT CAT CTT Leu Tyr Ser Leu Leu Thr Glu Gly Lys Arg Tyr Arg Ile Tyr His Leu 3000 3005 3010	9261
5	GCA ACT TCA AAA TCT AAA AGT AAA TCT GAA AGA GCT AAC ATA CAG TTA Ala Thr Ser Lys Ser Lys Ser Glu Arg Ala Asn Ile Gln Leu 3015 3020 3025	9309
10	GCA GCG ACA AAA AAA ACT CAG TAT CAA CAA CTA CCG GTT TCA GAT GAA Ala Ala Thr Lys Lys Thr Gln Tyr Gln Gln Leu Pro Val Ser Asp Glu 3030 3035 3040	9357
15	ATT TTA TTT CAG ATT TAC CAG CCA CGG GAG CCC CTT CAC TTC AGC AAA Ile Leu Phe Gln Ile Tyr Gln Pro Arg Glu Pro Leu His Phe Ser Lys 3045 3050 3055	9405
20	TTT TTA GAT CCA GAC TTT CAG CCA TCT TGT TCT GAG GTG GAC CTA ATA  Phe Leu Asp Pro Asp Phe Gln Pro Ser Cys Ser Glu Val Asp Leu Ile  3060 3065 3070 3075	9453
25	GGA TTT GTC GTT TCT GTT GTG AAA AAA ACA GGA CTT GCC CCT TTC GTC Gly Phe Val Val Ser Val Val Lys Lys Thr Gly Leu Ala Pro Phe Val 3080 3085 3090	9501
25	TAT TTG TCA GAC GAA TGT TAC AAT TTA CTG GCA ATA AAG TTT TGG ATA Tyr Leu Ser Asp Glu Cys Tyr Asn Leu Leu Ala Ile Lys Phe Trp Ile 3095 3100 3105	9549
30	GAC CTT AAT GAG GAC ATT ATT AAG CCT CAT ATG TTA ATT GCT GCA AGC Asp Leu Asn Glu Asp Ile Ile Lys Pro His Met Leu Ile Ala Ala Ser 3110 3120	9597
35	AAC CTC CAG TGG CGA CCA GAA TCC AAA TCA GGC CTT CTT ACT TTA TTT Asn Leu Gln Trp Arg Pro Glu Ser Lys Ser Gly Leu Leu Thr Leu Phe 3125 3130 3135	9645
40	GCT GGA GAT TTT TCT GTG TTT TCT GCT AGT CCA AAA GAG GGC CAC TTT Ala Gly Asp Phe Ser Val Phe Ser Ala Ser Pro Lys Glu Gly His Phe 3140 3145 3150 3155	9693
45	CAA GAG ACA TTC AAC AAA ATG AAA AAT ACT GTT GAG AAT ATT GAC ATA Gln Glu Thr Phe Asn Lys Met Lys Asn Thr Val Glu Asn Ile Asp Ile 3160 3165 3170	9741
43	CTT TGC AAT GAA GCA GAA AAC AAG CTT ATG CAT ATA CTG CAT GCA AAT Leu Cys Asn Glu Ala Glu Asn Lys Leu Met His Ile Leu His Ala Asn 3175 3180 3185	9789
50	GAT CCC AAG TGG TCC ACC CCA ACT AAA GAC TGT ACT TCA GGG CCG TAC Asp Pro Lys Trp Ser Thr Pro Thr Lys Asp Cys Thr Ser Gly Pro Tyr 3190 3195 3200	9837
55	ACT GCT CAA ATC ATT CCT GGT ACA GGA AAC AAG CTT CTG ATG TCT TCT Thr Ala Gln Ile Ile Pro Gly Thr Gly Asn Lys Leu Leu Met Ser Ser 3205 3210 3215	9885
60	CCT AAT TGT GAG ATA TAT TAT CAA AGT CCT TTA TCA CTT TGT ATG GCC Pro Asn Cys Glu Ile Tyr Tyr Gln Ser Pro Leu Ser Leu Cys Met Ala 3220 3225 3230 3235	9933
	AAA AGG AAG TCT GTT TCC ACA CCT GTC TCA GCC CAG ATG ACT TCA AAG	9981

	Lys Arg Lys Ser Val Ser Thr Pro Val Ser Ala Gln Met Thr Ser Lys 3240 3245 3250	
5	TCT TGT AAA GGG GAG AAA GAG ATT GAT GAC CAA AAG AAC TGC AAA AAG Ser Cys Lys Gly Glu Lys Glu Ile Asp Asp Gln Lys Asn Cys Lys Lys 3255 3260 3265	10029
10	AGA AGA GCC TTG GAT TTC TTG AGT AGA CTG CCT TTA CCT CCA CCT GTT Arg Arg Ala Leu Asp Phe Leu Ser Arg Leu Pro Leu Pro Pro Pro Val 3270 3280	10077
15	AGT CCC ATT TGT ACA TTT GTT TCT CCG GCT GCA CAG AAG GCA TTT CAG Ser Pro Ile Cys Thr Phe Val Ser Pro Ala Ala Gln Lys Ala Phe Gln 3285 3290 3295	10125
2.0	CCA CCA AGG AGT TGT GGC ACC AAA TAC GAA ACA CCC ATA AAG AAA AAA Pro Pro Arg Ser Cys Gly Thr Lys Tyr Glu Thr Pro Ile Lys Lys 3300 3305 3310 3315	10173
20	GAA CTG AAT TCT CCT CAG ATG ACT CCA TTT AAA AAA TTC AAT GAA ATT Glu Leu Asn Ser Pro Gln Met Thr Pro Phe Lys Lys Phe Asn Glu Ile 3320 3330	10221
25	TCT CTT TTG GAA AGT AAT TCA ATA GCT GAC GAA GAA CTT GCA TTG ATA Ser Leu Leu Glu Ser Asn Ser Ile Ala Asp Glu Glu Leu Ala Leu Ile 3335 3340 3345	10269
30	AAT ACC CAA GCT CTT TTG TCT GGT TCA ACA GGA GAA AAA CAA TTT ATA Asn Thr Gln Ala Leu Leu Ser Gly Ser Thr Gly Glu Lys Gln Phe Ile 3350 3360	10317
35	TCT GTC AGT GAA TCC ACT AGG ACT GCT CCC ACC AGT TCA GAA GAT TAT Ser Val Ser Glu Ser Thr Arg Thr Ala Pro Thr Ser Ser Glu Asp Tyr 3365 3370 3375	10365
40	CTC AGA CTG AAA CGA CGT TGT ACT ACA TCT CTG ATC AAA GAA CAG GAG Leu Arg Leu Lys Arg Arg Cys Thr Thr Ser Leu Ile Lys Glu Gln Glu 3380 3385 3390 3395	10413
40	AGT TCC CAG GCC AGT ACG GAA GAA TGT GAG AAA AAT AAG CAG GAC ACA Ser Ser Gln Ala Ser Thr Glu Glu Cys Glu Lys Asn Lys Gln Asp Thr 3400 3405 3410	10461
45	ATT ACA ACT AAA AAA TAT ATC TAA Ile Thr Thr Lys Lys Tyr Ile 3415	10485
50	(2) INFORMATION FOR SEQ ID NO:9:	
55	<ul><li>(i) SEQUENCE CHARACTERISTICS:</li><li>(A) LENGTH: 3418 amino acids</li><li>(B) TYPE: amino acid</li><li>(C) STRANDEDNESS: single</li><li>(D) TOPOLOGY: linear</li></ul>	
60	<ul><li>(ii) MOLECULE TYPE: protein</li><li>(v) FRAGMENT TYPE: internal</li><li>(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:</li></ul>	

	Met	Pro	Ile	Gly	Ser	Lys	Glu	Arg	Pro		Phe	Phe	Glu	Ile		Lys
	1 Thr	Arg	Cys	Asn	5 Lys	Ala	Asp	Leu	Gly	10 Pro	Ile	Ser	Leu		15 Trp	Phe
5	a1.	Glu	r	20	C 0.35	<i>α</i> 1	~ רת	Dro	25 Dro	Тугт	Λen	Ser	Glu	30 Pro	Δla	Glu
	GIU	GIU	ьеи 35	ser	ser	GIU	Ата	40	PIO	TYT	ASII	DCI	45	110	ALU	Olu
	Glu	Ser 50	Glu	His	Lys	Asn	Asn 55	Asn	Tyr	Glu	Pro	Asn 60	Leu	Phe	Lys	Thr
10	65	Gln				70					75					80
		Lys			85					90					95	
15		Leu	-	100					105					110		
		His	115					120					125			
	_	Val 130					135					140				
20	145	Leu				150					155					160
	_	Gly			165					170					175	
25		Lys		180					185					190		
		Trp	195					200					205			
		Ile 210					215					220				
30	225	Thr				230					235					240
	_	Lys			245					250					255	
35		Gln		260					265					270		
		Phe	275					280					285			
		Val 290					295					300				
40	305					310					315					320
		Lys			325					330					335	
45		Ala		340					345					350		
		Phe	355					360					365			
<b>5</b> 0		Val 370					375					380				
50	385					390					395					400
		Gly			405					410					415	
55		Ser		420					425					430		
		Lys	435					440					445			
60		Ser 450					455	•				460				
60	Val 465	Asn	. гуѕ	arg	Asp	470		. GIN	. nis	ьео	475		птр	1111	voh	480
		e Leu	Ala	Val	Lys	Gln	Ala	lle	Ser	Gly	Thr	Ser	Pro	Val	Ala	Ser

					485					490					495	
				500					505					510	Ser	
5	_		515					520					525		Pro	
		530					535					540			His	
10	545					550					555				Asp	560
	_				565					570					Lys 575	
				580					585					590	Ile	
15			595					600					605		Lys	
		610					615					620			Asn	
20	625					630					635				Leu	640
					645					650					Pro 655	
				660					665					670	Ser	
25			675					680					685		Asp	
	-	690		_	_		695					700			Thr	
30	705					710					715				Asn	720
		_			725					730					Ala 735	
				740					745					750	Thr	
35			755					760					765		Ser	
		770					775					780			Val	
40	785					790					795					Gly 800
					805					810					Met 815	
4.5				820					825					830	Val	
45			835					840					845		Arg	
		850					855					860				Gln Glu
50	865					870					875					880 Asn
					885					890					895	
55				900					905					910		Val
55	_		915	_				920					925			Lys
		930					935					940				Lys
60	945	_				950					955					960 Ser
				<u>,</u> -	965			1		970		•		-	975	

	τ	7	т1.	7 ~~	T	т1.	Dxo	Glu	Tara	7 an	7 cn	7 cn	The case	Mot	Λαn	Laze
	ьeu	ASII	iie	980	гуѕ	TIE	PIO		985	ASII	ASII	Asp	туг	990	Asp	цуз
	m	7.7.	<b>a</b> 1		T 0.11	<b>~1</b>	Dro	Ile		λαη	uic	Cor	Dhe		Gl v	Ser
_	Trp	Ата	_	ьeu	ьец	GIY	PIO	1000		ASII	nıs	ser	1005		Gry	Ser
5	-1	_	995		<b>a</b>	7	T			T	T	C			7	т1.
	Phe			Ala	Ser	Asn		Glu	ше	газ	ьeu			HIS	ASII	ıте
		1010		_			1015		_	3	~ 7	1020		_	<b>5</b>	m1
	-	-	Ser	Lys	Met			Lys	Asp	Ile			GIn	Tyr	Pro	
	1025					1030		_		_	1035		_	_	_	104
10	Ser	Leu	Ala	Cys	Val	Glu	Ile	Val	Asn			Ala	Leu	Asp		
					1045					1050				_	1055	
	Lys	Lys	Leu	Ser	Lys	Pro	Gln	Ser			Thr	Val	Ser	Ala	His	Leu
				1060					1065					1070		
	Gln	Ser	Ser	Val	Val	Val	Ser	Asp	Cys	Lys	Asn	Ser	His	Ile	Thr	Pro
15			1075					1080					1085			
	Gln	Met	Leu	Phe	Ser	Lys	Gln	Asp	Phe	Asn	Ser	Asn	His	Asn	Leu	Thr
		1090					1099					1100				
	Pro	Ser	Gln	Lys	Ala	Glu	Ile	Thr	Glu	Leu	Ser	Thr	Ile	Leu	Glu	Glu
	1105	5				1110	)				1115	5				112
20	Ser	Gly	Ser	Gln	Phe	Glu	Phe	Thr	Gln	Phe	Arg	Lys	Pro	Ser	Tyr	Ile
		_			1125	5				1130	)				1135	5
	Leu	Gln	Lys	Ser	Thr	Phe	Glu	Val	Pro	Glu	Asn	Gln	Met	Thr	Ile	Leu
			_	1140	)				1145	5				1150	)	
	Lys	Thr	Thr	Ser	Glu	Glu	Cys	Arg	Asp	Ala	Asp	Leu	His	Val	Ile	Met
25	-		1155				-	1160					1165			
	Asn	Ala	Pro	Ser	Ile	Gly	Gln	Val	Asp	Ser	Ser	Lys	Gln	Phe	Glu	Gly
		1170				-	1175		-			1180				_
	Thr			Ile	Lvs	Ara	Lvs	Phe	Ala	Gly	Leu	Leu	Lys	Asn	Asp	Cys
	118				1	1190				•	1199		-		-	120
30	Asn	Lvs	Ser	Ala	Ser			Leu	Thr	Asp	Glu	Asn	Glu	Val	Gly	Phe
		-1-			1205		2			1210					121!	
	Ara	Glv	Phe	Tyr			His	Gly	Thr	Lys	Leu	Asn	Val	Ser	Thr	Glu
		1		1220				-	122					1230		
	Ala	Leu	Gln	Lvs	Ala	Val	Lys	Leu	Phe	Ser	Asp	Ile	Glu	Asn	Ile	Ser
35			123				-	1240			-		124			
	Glu	Glu			Ala	Glu	Val	His	Pro	Ile	Ser	Leu	Ser	Ser	Ser	Lys
		1250					125					1260				-
	Cvs			Ser	Val	Val	Ser	Met	Phe	Lys	Ile	Glu	Asn	His	Asn	Asp
	126					1270				-	127					128
40			Val	Ser	Glu	Lys	Asn	Asn	Lys	Cys	Gln	Leu	Ile	Leu	Gln	Asn
	4				1289				•	129					129	
	Asn	Ile	Glu	Met	Thr	Thr	Glv	Thr	Phe	Val	Glu	Glu	Ile	Thr	Glu	Asn
	Tvr															Ser
45	1		131					1320			•	-	132			
	Ara	Asn			Asn	Leu	Glu			Glv	Ser	asp	Ser	Ser	Lys	Asn
	5	133					133					134			_	
	Asp			Cvs	Ile	His			Glu	Thr	Asp	Leu	Leu	Phe	Thr	Asp
	134			-1-		1350					135					136
50			Asn	Ile	Cvs			Leu	Ser	Glv			Met	Lys	Glu	Gly
	0				136		2			137				1	137	
	Asn	Thr	Gln	Ile			Asp	Leu	Ser			Thr	Phe	Leu	Glu	Val
				138					138					139		
	Ala	Lvs	Ala			Ala	Cvs	His			Thr	Ser	Asn	Lvs	Glu	Gln
55		-7-	139				-1-	140					140			
33	Len	Thr			Lvs	Thr	Glu			Ile	Lvs	Asp			Thr	Ser
	пси	141			<b>L</b> , 5	1111	141					142				
	Acn			Phe	Gln	Thr			Glv	Lvs	Asn			Val	Ala	Lys
	142		1116	1110	OIII	143		UCL	O ± Y	y 5	143		201			144
60			Phe	Aan	Laze			Agn	Phe	Phe			Lvs	Pro	Glu	Glu
50	Jiu		- 110		144			. 1,011		145			-10		145	
	יים. ן	Hie	Δen	Phe			Agn	Ser	Glu			Ser	Agn	Ile		Lys
	u	****	* ******	تبا د د د	~	u			4	u					5	

		1460	1.	465	1470
	147	Asp Ile Leu 5	1480		Ile Val Lys His 1485
5	Lys Ile Leu 1490	Lys Glu Ser	Val Pro V 1495	al Gly Thr Gly 150	r Asn Gln Leu Val
	Thr Phe Gln 1505	Gly Gln Pro		sp Glu Lys Ile 1515	Lys Glu Pro Thr 152
10	Leu Leu Gly	Phe His Thr 1525	Ala Ser G	ly Lys Lys Val 1530	Lys Ile Ala Lys 1535
	Glu Ser Leu	Asp Lys Val		eu Phe Asp Glu 545	Lys Glu Gln Gly 1550
	Thr Ser Glu	Ile Thr Ser	Phe Ser H 1560	is Gln Trp Ala	Lys Thr Leu Lys 1565
15	Tyr Arg Glu 1570	Ala Cys Lys	Asp Leu G 1575	lu Leu Ala Cys 158	Glu Thr Ile Glu
	Ile Thr Ala	Ala Pro Lys		lu Met Gln Asr 1595	Ser Leu Asn Asn 160
2.0			Ile Glu T		Pro Lys Leu Leu
20	Ser Asp Asn	1605 Leu Cys Arg 1620		1610 Ilu Asn Leu Lys 625	1615 Thr Ser Lys Ser 1630
	Ile Phe Leu 163	Lys Val Lys			Lys Glu Thr Ala 1645
25				sn Gln Ser Pro	Tyr Ser Val Ile
		Ala Leu Ala 167	Phe Tyr T		Arg Lys Thr Ser
30					Leu Arg Glu Gly 1695
<b>J</b> •	Ile Phe Asp		_		Asp Tyr Val Gly 1710
	Asn Tyr Leu 171	Tyr Glu Asn			e Ala Glu Asn Asp 1725
35	Lys Asn His	Leu Ser Glu		sp Thr Tyr Let 174	Ser Asn Ser Ser
	Met Ser Asn 1745	Ser Tyr Ser 175		er Asp Glu Val 1755	l Tyr Asn Asp Ser 176
40	Gly Tyr Leu	Ser Lys Asn 1765	Lys Leu A	sp Ser Gly Ile 1770	e Glu Pro Val Leu 1775
	Lys Asn Val	Glu Asp Gln 1780		thr Ser Phe Ser .785	Lys Val Ile Ser 1790
	Asn Val Lys 179	_	Ala Tyr P 1800	Pro Gln Thr Val	l Asn Glu Asp Ile 1805
45	1810		1815	182	
	1825	183	0	1835	n Phe Glu Val Gly 184
50	Pro Pro Ala	Phe Arg Ile 1845	Ala Ser G	ly Lys Ile Val 1850	l Cys Val Ser His 1855
		1860	1	.865	Ser Phe Ser Lys 1870
	187	5	1880		e Cys Gln Thr Lys 1885
55	1890		1895	190	
	His Asn Ser 1905	Leu Asp Asn 191		Cys Ser Thr His 1915	s Ser His Lys Val 192
60	Phe Ala Asp	Ile Gln Ser 1925	Glu Glu I	le Leu Gln His	s Asn Gln Asn Met 1935
	Ser Gly Leu	Glu Lys Val 1940		lle Ser Pro Cys .945	s Asp Val Ser Leu 1950

	Glu	Thr	Ser 1955		Ile	Cys	Lys	Cys 1960		Ile	Gly	Lys	Leu 1965		Lys	Ser
5	Val	Ser 1970		Ala	Asn		Cys 1975		Ile	Phe		Thr 1980		Ser	Gly	Lys
	Ser 1985	Val		Val	Ser	Asp 1990		Ser	Leu	Gln	Asn 1995		Arg	Gln	Val	Phe 200
	Ser	Glu	Ile	Glu	Asp 2005		Thr	Lys	Gln	Val 2010		Ser	Lys	Val	Leu 2015	
10	_			2020	)				2025	5				2030		
			2035	5				2040	)				2045	5	Tyr	
15		2050	)				2055	5				2060	)		Gly	
	2065	5				2070	)				2075	5			Val	208
					2085	5				2090	)				Ser 2095	5
20			_	2100	0				210	5				211		
			2115	5				2120	)				212	5	Ser	
25		213	0				2135	5				214	0		Ser Gln	
	2145	5				2150	0				215	5			Glu	216
2.0	-	-			216	5				217	0				2179 Lys	5
30				218	0				218	5				219		
			219	5				2200	C				220	5	Phe	
35		221	0				221	5				222	0		Glu	
	222	5				223	0				223	5			Thr	224
40					224	5			Ser	225 Asn	0			Gly	225! Lys	5
					Leu					Glu					Arg	
45			Asn					Ile					Glu		Ser	
40	Lys 230	Ala		Lys	Ser	Thr 231	Pro		Gly	Thr	Ile 231	Lys		Arg	Arg	Let 232
			His	His	Val 232	Ser		Glu	Pro	Ile 233	Thr		Val	Pro	Phe 233	Arg
50	Thr	Thr	Lys	Glu 234	Arg		Glu	Ile	Gln 234		Pro	Asn	Phe	Thr 235	Ala 0	Pro
	Gly	Gln	Glu 235		Leu	Ser	Lys	Ser 236		Leu	Tyr	Glu	His 236		Thr	Leu
55		237	0				237	5				238	0		Tyr	
	238	5				239	0				239	5			Thr	240
					240	5				241	0				His 241	5
60		_		242	0	_		_	242	5				243		
	(2 l m	LAVE	(3 l n	Acn		Agn	(4137	HIG	(413)	Ser	Agn	ASD	ser	LVS	Asn	LiV

		2435	5				2440					2445			
	Ile Asn 245	Asp		Glu	Ile		Gln		Asn		Asn 2460		Ser	Asn	Gln
5	Ala Ala 2465	Ala	Val		Phe 2470		Lys	Cys		Glu 2475		Pro	Leu	Asp	Leu 248
	Ile Thr	Ser	Leu		Asn		Arg	Asp	Ile 2490		Asp	Met	Arg	Ile 2495	
10	Lys Lys	Gln	Arg 2500	Gln		Val		Pro 2505	Gln		Gly	Ser	Leu 2510		Leu
_ ~	Ala Lys	Thr 2519	Ser		Leu			Ile		Leu	Lys	Ala 2525		Val	Gly
	Gly Glr 253	Val		Ser	Ala		Ser		Lys	Gln	Leu 2540		Thr	Tyr	Gly
15	Val Ser 2545		His		Ile 2550	Lys		Asn	Ser	Lys 2555	Asn		Glu	Ser	Phe 256
	Gln Phe	His	Thr		Asp		Phe	Gly	Lys 2570	Glu		Leu	Trp	Thr 2575	Gly
20	Lys Gly	Ile	Gln 2580	Leu		Asp		Gly 2585	Trp		Ile	Pro	Ser 2590	Asn	
20	Gly Lys	Ala 2599	Gly		Glu	Glu		Tyr		Ala	Leu	Cys 2605	Asp		Pro
	Gly Val	Asp		Lys	Leu	Ile 2615	Ser		Ile	Trp	Val 2620	Tyr		His	Tyr
25	Arg Trp		Ile	Trp	Lys 2630	Leu		Ala	Met	Glu 2635		Ala	Phe	Pro	Lys 264
	Glu Phe	Ala	Asn	Arg 2645	_	Leu	Ser	Pro	Glu 2650		Val	Leu	Leu	Gln 2655	
30	Lys Tyr	Arg	Tyr 2660	Asp		Glu	Ile	Asp 2665		Ser	Arg	Arg	Ser 2670		Ile
	Lys Lys	Ile 267!	Met		Arg	Asp	Asp 2680		Ala	Ala	Lys	Thr 2685		Val	Leu
	Cys Val	Ser		Ile	Ile	Ser 2695		Ser	Ala	Asn	Ile 2700		Glu	Thr	Ser
35	Ser Asi 2705	ı Lys	Thr	Ser	Ser 2710		Asp	Thr	Gln	Lys 2715		Ala	Ile	Ile	Glu 272
	Leu Thi			2725	5				2730	)				2735	5
40	Leu Ala		2740	)				2745	5				2750	)	
	Leu His	275	5				2760	)				2765	5		
	Glu Ala 27	70				2775	5				278	)			
45	Pro Ala 2785	_	-	-	2790	) _				2795	5				280
	Phe Pro			2809	5				2810	)				281	5
50	Cys Va	_	2820	)				282	5				2830	)	
	Lys Th	283	5	_		_	2840	)				284	5		
	Lys Glu 28	50		-	-	2855	5				286	0			
55	Leu Pho				2870	)				287	5				288
	Thr Ly			288	5				289	С				289	5
60	Ala Le		2900	)				290	5				291	0	
	Asp Pr	o Ala 291	_	ьeu	Glu	GLY	Tyr 292		ser	Glu	GIU	GIn 292		arg	Ala

	Leu	Asn 2930		His	Arg	Gln	Met 2935		Asn	Asp	Lys	Lys 2940	Gln	Ala	Gln	Ile
	Gln	Leu	Glu	Ile	Ara	Lvs	Ala	Met	Glu	Ser	Ala	Glu	Gln	Lys	Glu	Gln
5	2945		014		5	2950					2955			-		296
J	$G_{1}^{2}$	, T.211	Ser	Δrα	Asn			Thr	Val	Trn			Ara	Ile	Val	Ser
	Gry	пси	DCI	nr 9	2965					2970					2975	
	TT 2.25	cor	Lys	Tare			Agn	Ser				Ser	Tle			
	1 7 1	per	пуъ	2980		цуз	ASP	DCI	2985		200	001		2990	J	
10	Ser	Ser	Asp			Ser			Thr		Gly	Lys		Tyr		Ile
			2995					3000					3005		_	
	_	3010	Leu				3015	5				3020	)			
	Ile	Gln	Leu	Ala	Ala	Thr	Lys	Lys	Thr	Gln	Tyr	Gln	Gln	Leu	Pro	Val
15	302					3030		_			3035	5				304
	Ser	Asp	Glu	Ile	Leu	Phe	Gln	Ile	Tyr	Gln	Pro	Arg	Glu	Pro	Leu	His
		<u>F</u>			3045				_	3050					3055	
	Phe	Ser	Lys	Phe	Leu	Asp	Pro	Asp	Phe	Gln	Pro	Ser	Cys	Ser	Glu	Val
				3060				-	3065				-	3070		
20	Asp	Leu	Ile	Gly		Val	Val	Ser 3080	Val		Lys	Lys	Thr 3089		Leu	Ala
	D	Db -	3075 Val		T 011	Cor	7 an			Tur	λen	Len			Tle	Lvs
	Pro			Tyr	ьеи	ser	3099		Cys	TAT	ASII	310	n Deu	AΙα	110	цуь
	<b>51</b>	309	Ile	3	r	7.00			Tla	Tla	Lare	-		Met	T.e.ii	Tle
0.5			шe	Asp	ьeu			Asp	116	116	311!		1110	PIC C	пси	312
25	310	5 - 3	_	_	<b>.</b>	311		7	Dago	a1			Cor	Gly	T.011	
	Ala	Ala	Ser	Asn			Trp	Arg	PIO			пур	SEI	GIY	313!	
	_				312				77-7	313		77-	C ~ ~	Dec		
	Thr	Leu	Phe			Asp	Phe	Ser			Ser	Ala	ser			Giu
				314		_			314!		_		m1	3150		7. ~ ~
30	Gly	His	Phe	Gln	Glu	Thr	Phe			Met	Lys	Asn			Glu	Asn
			315				_	3160		_	_	_	316		<b>-1</b> .	<b>.</b>
	Ile	Asp	Ile	Leu	Cys	Asn			Glu	Asn	Lys			HIS	тте	ьeu
		317					317		_			318		_	1	
	His	Ala	Asn	Asp	Pro			Ser	Thr	Pro			Asp	Cys	Inr	
35	318	5				319					319		_	_	_	320
	Gly	Pro	Tyr	Thr	Ala	Gln	Ile	Ile	Pro	Gly	Thr	Gly	Asn	Lys		
					320	5				321			_	_	321	
	Met	Ser	Ser	Pro	Asn	Cys	Glu	Ile			Gln	Ser	Pro	Leu	Ser	Leu
				322					322					323		
40	Cys	Met	Ala	Lys	Arg	Lys	Ser	Val	Ser	Thr	Pro	Val			Gln	Met
			323					324					324			
	Thr	Ser	Lys	Ser	Cys	Lys	Gly	Glu	Lys	Glu	Ile	Asp	Asp	Gln	Lys	Asn
		325	0				325	5				326	0			
	Cys	Lys	Lys	Arg	Arg	Ala	Leu	Asp	Phe	Leu	Ser	Arg	Leu	Pro	Leu	
45	326	5				327					327					328
	Pro	Pro	Val	Ser	Pro	Ile	Cys	Thr	Phe	Val	Ser	Pro	Ala	Ala	Gln	Lys
					328	5				329	0				329	5
	Ala	Phe	Gln	Pro	Pro	Arg	ser,	Cys	Gly	Thr	Lys	Tyr	Glu	Thr	Pro	Ile
				330	0				330	5				331	0	
50	Lys	Lys	Lys	Glu	Leu	Asn	Ser	Pro	Gln	Met	Thr	Pro	Phe	Lys	Lys	Phe
	•	•	331					332					332			
	Asr	Glu	ı Ile	Ser	Leu	Leu	ı Glu	Ser	Asn	Ser	Ile	Ala	Asp	Glu	Glu	Leu
		333					333					334				
	Ala	Lei	ı Ile	Asn	Thr	Glr			Leu	Ser	Gly	Ser	Thr	Gly	Glu	Lys
55	334					335					335			-		336
<b>J J</b>			e Ile	Ser	· Val			Ser	Thr	Ara	Thr	· Ala	Pro	Thr	Ser	Ser
	J_1				336		<del>-</del>		<b>-</b>	337					337	
	G1,	ı Aer	Tyr	· J.e.			ı Lıvs	Ara	Ara			Thr	Ser	Leu		
	310		- <u>- 7</u> -	338		,		5	338					339		-
60	ر 1	1 (31 r	ı Glu			Glr	ı Ala	Ser			Glu	CVS	Glu			Lys
0.0	010		339					340			•	1 -	340			_
	<b>a</b> 1	. 7			mb-	- mb-	c T 1/0			т 1 с						

3410 3415

	5			(2)	INF	ORMA	TION	FOR	SEQ	ID	NO:1	0:						
	10		(i	(A) (B) (C)	LENG TYPE STRA	TH: : nu NDED	1048 clei NESS	CTER 5 ba c ac : do near	se p id uble	airs								
			•		IOLEC	ULE		: cD										
	15			(B)	LOC	ATIC	N: 2	odin 29 MATI	.104	82		MI4)						
	20		( x	ai) S	EQUE	NCE	DESC	RIPT	'ION:	SEÇ	ID	NO : 1	.0:					
į		TCTC	GCTGC GATTI	CGC C	TCGG	GTG1	C TI	TTGC TTTT	GGCG GTCA	GTG GCI	GGTC TACT	GCC CCG	GCCA	GGAC AAAA	AA G AG A ATG	CGTG ACTG CCT	GCGCC SAGGGG CACCT ATT	60 120 180 237
	25														Met 1	: Prc	lle	
	30		TCC Ser 5															285
	35	AAC Asn 20	AAA Lys	GCA Ala	GAT Asp	TTA Leu	GGA Gly 25	CCA Pro	ATA Ile	AGT Ser	CTT Leu	AAT Asn 30	TGG Trp	TTT Phe	GAA Glu	GAA Glu	CTT Leu 35	333
	33		TCA Ser															381
	40		AAA Lys															429
	45	AAA Lys	CCA Pro	TCT Ser 70	TAT Tyr	AAT Asn	CAG Gln	CTG Leu	GCT Ala 75	TCA Ser	ACT Thr	CCA Pro	ATA Ile	ATA Ile 80	TTC Phe	AAA Lys	GAG Glu	477
	50		GGG Gly 85															525
			TTC Phe															573
	55		CTT Leu															621
	60		CCA Pro															669

5	TGT Cys	ACA Thr	CAT His 150	GTA Val	ACA Thr	CCA Pro	CAA Gln	AGA Arg 155	GAT Asp	AAG Lys	TCA Ser	GTG Val	GTA Val 160	TGT Cys	GGG Gly	AGT Ser	717
10	TTG Leu	TTT Phe 165	CAT His	ACA Thr	CCA Pro	AAG Lys	TTT Phe 170	GTG Val	AAG Lys	GGT Gly	CGT Arg	CAG Gln 175	ACA Thr	CCA Pro	AAA Lys	CAT His	765
10	ATT Ile 180	TCT Ser	GAA Glu	AGT Ser	CTA Leu	GGA Gly 185	GCT Ala	GAG Glu	GTG Val	GAT Asp	CCT Pro 190	GAT Asp	ATG Met	TCT Ser	TGG Trp	TCA Ser 195	813
15	AGT Ser	TCT Ser	TTA Leu	GCT Ala	ACA Thr 200	CCA Pro	CCC Pro	ACC Thr	CTT Leu	AGT Ser 205	TCT Ser	ACT Thr	GTG Val	CTC Leu	ATA Ile 210	GTC Val	861
20	AGA Arg	AAT Asn	GAA Glu	GAA Glu 215	GCA Ala	TCT Ser	GAA Glu	ACT Thr	GTA Val 220	TTT Phe	CCT Pro	CAT His	GAT Asp	ACT Thr 225	ACT Thr	GCT Ala	909
25													CTG Leu 240				957
20	GAT Asp	AGA Arg 245	TTT Phe	ATC Ile	GCT Ala	TCT Ser	GTG Val 250	ACA Thr	GAC Asp	AGT Ser	GAA Glu	AAC Asn 255	ACA Thr	AAT Asn	CAA Gln	AGA Arg	1005
30													AAT Asn				1053
35													CCA Pro				1101
40													GAA Glu		Asp		1149
45				Cys									CTA Leu 320			GTA Val	1197
50	AGA Arg	ACT Thr 325	Ser	AAG Lys	ACT Thr	AGG Arg	AAA Lys 330	Lys	ATT Ile	TTC Phe	CAT His	GAA Glu 335	GCA Ala	AAC Asn	GCT Ala	GAT Asp	1245
50		Cys					Asn					Lys	TAC Tyr			GTA Val 355	1293
55						Asn					Leu					GCA Ala	1341
60					Phe					Asp					Glu	GTT Val	1389

r			TCT Ser 390														1437
5			GCC Ala														1485
10			AAT Asn														1533
15			GAT Asp														1581
20			AAA Lys														1629
25			GAA Glu 470														1677
23			CAG Gln														1725
30			AAA Lys														1773
35			GCA Ala														1821
40			GAA Glu														1869
45			GAG Glu 550														1917
			ACC Thr												_		1965
50		Ser	ACT Thr														2013
55			ACA Thr														2061
60			ATT							Glu						GCA Ala	2109
	CCA	CTT	ACA	TTT	GCA	AAT	GCT	GAT	TCA	GGT	TTA	TTG	CAT	TCT	TCT	GTG	2157

	Pro	Leu	Thr 630	Phe	Ala	Asn	Ala	Asp 635	Ser	Gly	Leu	Leu	His 640	Ser	Ser	Val	
5				TGT Cys													2205
10				TTT Phe													2253
15				AAT Asn													2301
20	AAA Lys	TGT Cys	AAT Asn	AAG Lys 695	GAA Glu	AAA Lys	CTA Leu	CAG Gln	TTA Leu 700	TTT Phe	ATT Ile	ACC Thr	CCA Pro	GAA Glu 705	GCT Ala	GAT Asp	2349
20				TGC Cys													2397
25				TCA Ser													2445
30	CCA Pro 740	GTA Val	CAA Gln	CAT His	TCA Ser	AAA Lys 745	GTG Val	GAA Glu	TAC Tyr	AGT Ser	GAT Asp 750	ACT Thr	GAC Asp	TTT Phe	CAA Gln	TCC Ser 755	2493
35				CTT Leu													2541
4.0				TCC Ser 775												AGA Arg	2589
40				TCA Ser												TAT Tyr	2637
45			Asp													CAA Gln	2685
50		Val					Glu					. Val				CCA Pro 835	2733
55						Arg					Ser					TTC Phe	2781
60					Asn					Gln					Glu	ACT Thr	2829
60																TTC Phe	2877

870 875 880

5	TCA Ser	GAC Asp 885	AAT Asn	GAG Glu	AAT Asn	AAT Asn	TTT Phe 890	GTC Val	TTC Phe	CAA Gln	GTA Val	GCT Ala 895	AAT Asn	GAA Glu	AGG Arg	AAT Asn	2925
10	AAT Asn 900	CTT Leu	GCT Ala	TTA Leu	GGA Gly	AAT Asn 905	ACT Thr	AAG Lys	GAA Glu	CTT Leu	CAT His 910	GAA Glu	ACA Thr	GAC Asp	TTG Leu	ACT Thr 915	2973
				GAA Glu													3021
15				GAT Asp 935													3069
20				CTT Leu													3117
25	AAA Lys	ATG Met 965	ACT Thr	CTA Leu	GGT Gly	CAA Gln	GAT Asp 970	TTA Leu	AAA Lys	TCG Ser	GAC Asp	ATC Ile 975	TCC Ser	TTG Leu	AAT Asn	ATA Ile	3165
30				CCA Pro													3213
				CCA Pro					Ser					Phe			3261
35			Asn	AAG Lys 1015				Leu					Ile				3309
40		Met		TTC Phe			Ile					Pro					3357
45	Cys			ATT Ile		Asn					Asp						3405
50		Lys		CAG Gln	Ser		Asn			Ser		His			Ser	AGT Ser 1075	3453
				Ser		Cys			Ser		Ile						3501
55					Asp			Ser							Ser	CAA Gln	3549
60				Ile					Thr					Ser		AGT Ser	3597

5	CAG TTT GAA TTT ACT CAG TTT AGA AAG CCA AGC TAC ATA TTG CAG AAG Gln Phe Glu Phe Thr Gln Phe Arg Lys Pro Ser Tyr Ile Leu Gln Lys 1125 1130 1135	3645
10	AGT ACA TTT GAA GTG CCT GAA AAC CAG ATG ACT ATC TTA AAG ACC ACT Ser Thr Phe Glu Val Pro Glu Asn Gln Met Thr Ile Leu Lys Thr Thr 1140 1145 1150 1155	3693
10	TCT GAG GAA TGC AGA GAT GCT GAT CTT CAT GTC ATA ATG AAT GCC CCA Ser Glu Glu Cys Arg Asp Ala Asp Leu His Val Ile Met Asn Ala Pro 1160 1165 1170	3741
15	TCG ATT GGT CAG GTA GAC AGC AGC AAG CAA TTT GAA GGT ACA GTT GAA Ser Ile Gly Gln Val Asp Ser Ser Lys Gln Phe Glu Gly Thr Val Glu 1175 1180 1185	3789
20	ATT AAA CGG AAG TTT GCT GGC CTG TTG AAA AAT GAC TGT AAC AAA AGT Ile Lys Arg Lys Phe Ala Gly Leu Leu Lys Asn Asp Cys Asn Lys Ser 1190 1195 1200	3837
25	GCT TCT GGT TAT TTA ACA GAT GAA AAT GAA GTG GGG TTT AGG GGC TTT Ala Ser Gly Tyr Leu Thr Asp Glu Asn Glu Val Gly Phe Arg Gly Phe 1205 1210 1215	3885
2.0	TAT TCT GCT CAT GGC ACA AAA CTG AAT GTT TCT ACT GAA GCT CTG CAA Tyr Ser Ala His Gly Thr Lys Leu Asn Val Ser Thr Glu Ala Leu Gln 1220 1235	3933
30	AAA GCT GTG AAA CTG TTT AGT GAT ATT GAG AAT ATT AGT GAG GAA ACT Lys Ala Val Lys Leu Phe Ser Asp Ile Glu Asn Ile Ser Glu Glu Thr 1240 1245 1250	3981
35	TCT GCA GAG GTA CAT CCA ATA AGT TTA TCT TCA AGT AAA TGT CAT GAT Ser Ala Glu Val His Pro Ile Ser Leu Ser Ser Lys Cys His Asp 1255 1260 1265	4029
40	TCT GTT GTT TCA ATG TTT AAG ATA GAA AAT CAT AAT GAT AAA ACT GTA Ser Val Val Ser Met Phe Lys Ile Glu Asn His Asn Asp Lys Thr Val 1270 1275 1280	4077
45	AGT GAA AAA AAT AAA TGC CAA CTG ATA TTA CAA AAT AAT ATT GAA Ser Glu Lys Asn Asn Lys Cys Gln Leu Ile Leu Gln Asn Asn Ile Glu 1285 1290 1295	4125
50	ATG ACT ACT GGC ACT TTT GTT GAA GAA ATT ACT GAA AAT TAC AAG AGA Met Thr Thr Gly Thr Phe Val Glu Glu Ile Thr Glu Asn Tyr Lys Arg 1300 1305 1310 1315	4173
30	AAT ACT GAA AAT GAA GAT AAC AAA TAT ACT GCT GCC AGT AGA AAT TCT Asn Thr Glu Asn Glu Asp Asn Lys Tyr Thr Ala Ala Ser Arg Asn Ser 1320 1325 1330	4221
55	CAT AAC TTA GAA TTT GAT GGC AGT GAT TCA AGT AAA AAT GAT ACT GTT His Asn Leu Glu Phe Asp Gly Ser Asp Ser Ser Lys Asn Asp Thr Val 1335 1340 1345	4269
60	TGT ATT CAT AAA GAT GAA ACG GAC TTG CTA TTT ACT GAT CAG CAC AAC Cys Ile His Lys Asp Glu Thr Asp Leu Leu Phe Thr Asp Gln His Asn 1350 1355 1360	4317

5		CTT AAA TTA Leu Lys Leu			Met Lys G			j
3		GAA GAT TTG Glu Asp Leu						,
10		GCA TGT CAT Ala Cys His 1400		Thr Ser		Slu Gln Leu		-
15		ACG GAG CAA Thr Glu Gln 1415						<b>)</b>
20	Phe Gln 7	ACT GCA AGT Thr Ala Ser 430	Gly Lys					7
25		ATT GTA AAT Ile Val Asn			Lys Pro G			;
		TTA AAT TCT Leu Asn Ser						}
30		CTA AGT TAT Leu Ser Tyr 1480		Thr Asp		ys His Lys		L
35		AGT GTC CCA Ser Val Pro 1495						)
40	Gly Gln	CCC GAA CGT Pro Glu Arg 510	Asp Glu					7
45		ACA GCT AGC Thr Ala Ser			Lys Ile A			5
		GTG AAA AAC Val Lys Asn						3
50		AGT TTT AGC Ser Phe Ser 1560		Trp Ala		Leu Lys Tyr		L
55		AAA GAC CTT Lys Asp Leu 1575						Э
60	Ala Pro	AAG TGT AAA Lys Cys Lys 590	Glu Met					7
	CTT GTT	TCT ATT GAG	ACT GTG	GTG CCA	CCT AAG C	CTC TTA AGT	GAT AAT 508	5

	Leu Val Ser Ile Glu Thr Val Val Pro Pro Lys Leu Leu Ser Asp Asn 1605 1610 1615	
5	TTA TGT AGA CAA ACT GAA AAT CTC AAA ACA TCA AAA AGT ATC TTT TTG Leu Cys Arg Gln Thr Glu Asn Leu Lys Thr Ser Lys Ser Ile Phe Leu 1620 1625 1630 1635	5133
10	AAA GTT AAA GTA CAT GAA AAT GTA GAA AAA GAA ACA GCA AAA AGT CCT Lys Val Lys Val His Glu Asn Val Glu Lys Glu Thr Ala Lys Ser Pro 1640 1645 1650	5181
15	GCA ACT TGT TAC ACA AAT CAG TCC CCT TAT TCA GTC ATT GAA AAT TCA Ala Thr Cys Tyr Thr Asn Gln Ser Pro Tyr Ser Val Ile Glu Asn Ser 1655 1660 1665	5229
20	GCC TTA GCT TTT TAC ACA AGT TGT AGT AGA AAA ACT TCT GTG AGT CAG Ala Leu Ala Phe Tyr Thr Ser Cys Ser Arg Lys Thr Ser Val Ser Gln 1670 1675 1680	5277
20	ACT TCA TTA CTT GAA GCA AAA AAA TGG CTT AGA GAA GGA ATA TTT GAT Thr Ser Leu Leu Glu Ala Lys Lys Trp Leu Arg Glu Gly Ile Phe Asp 1685 1690 1695	5325
25	GGT CAA CCA GAA AGA ATA AAT ACT GCA GAT TAT GTA GGA AAT TAT TTG Gly Gln Pro Glu Arg Ile Asn Thr Ala Asp Tyr Val Gly Asn Tyr Leu 1700 1705 1710 1715	5373
30	TAT GAA AAT AAT TCA AAC AGT ACT ATA GCT GAA AAT GAC AAA AAT CAT Tyr Glu Asn Asn Ser Asn Ser Thr Ile Ala Glu Asn Asp Lys Asn His 1720 1725 1730	5421
35	CTC TCC GAA AAA CAA GAT ACT TAT TTA AGT AAC AGT AGC ATG TCT AAC Leu Ser Glu Lys Gln Asp Thr Tyr Leu Ser Asn Ser Ser Met Ser Asn 1735 1740 1745	5469
40	AGC TAT TCC TAC CAT TCT GAT GAG GTA TAT AAT GAT TCA GGA TAT CTC Ser Tyr Ser Tyr His Ser Asp Glu Val Tyr Asn Asp Ser Gly Tyr Leu 1750 1760	5517
10	TCA AAA AAT AAA CTT GAT TCT GGT ATT GAG CCA GTA TTG AAG AAT GTT Ser Lys Asn Lys Leu Asp Ser Gly Ile Glu Pro Val Leu Lys Asn Val 1765 1770 1775	5565
45	GAA GAT CAA AAA AAC ACT AGT TTT TCC AAA GTA ATA TCC AAT GTA AAA Glu Asp Gln Lys Asn Thr Ser Phe Ser Lys Val Ile Ser Asn Val Lys 1780 1785 1790 1795	5613
50	GAT GCA AAT GCA TAC CCA CAA ACT GTA AAT GAA GAT ATT TGC GTT GAG Asp Ala Asn Ala Tyr Pro Gln Thr Val Asn Glu Asp Ile Cys Val Glu 1800 1805 1810	5661
55	GAA CTT GTG ACT AGC TCT TCA CCC TGC AAA AAT AAA AAT GCA GCC ATT Glu Leu Val Thr Ser Ser Pro Cys Lys Asn Lys Asn Ala Ala Ile 1815 1820 1825	5709
60	AAA TTG TCC ATA TCT AAT AGT AAT AAT TTT GAG GTA GGG CCA CCT GCA Lys Leu Ser Ile Ser Asn Ser Asn Asn Phe Glu Val Gly Pro Pro Ala 1830 1835 1840	5757
00	TTT AGG ATA GCC AGT GGT AAA ATC GTT TGT GTT TCA CAT GAA ACA ATT Phe Arg Ile Ala Ser Gly Lys Ile Val Cys Val Ser His Glu Thr Ile	5805

1845 1850 1855

5	AAA AAA GTG AAA GAC ATA TTT ACA GAC AGT TTC AGT AAA GTA ATT AAG Lys Lys Val Lys Asp Ile Phe Thr Asp Ser Phe Ser Lys Val Ile Lys 1860 1865 1870 1875	5853
10	GAA AAC AAC GAG AAT AAA TCA AAA ATT TGC CAA ACG AAA ATT ATG GCA Glu Asn Asn Glu Asn Lys Ser Lys Ile Cys Gln Thr Lys Ile Met Ala 1880 1885 1890	5901
15	GGT TGT TAC GAG GCA TTG GAT GAT TCA GAG GAT ATT CTT CAT AAC TCT Gly Cys Tyr Glu Ala Leu Asp Asp Ser Glu Asp Ile Leu His Asn Ser 1895 1900 1905	5949
	CTA GAT AAT GAT GAA TGT AGC ACG CAT TCA CAT AAG GTT TTT GCT GAC Leu Asp Asn Asp Glu Cys Ser Thr His Ser His Lys Val Phe Ala Asp 1910 1915 1920	5997
20	ATT CAG AGT GAA GAA ATT TTA CAA CAT AAC CAA AAT ATG TCT GGA TTG Ile Gln Ser Glu Glu Ile Leu Gln His Asn Gln Asn Met Ser Gly Leu 1925 1930 1935	6045
25	GAG AAA GTT TCT AAA ATA TCA CCT TGT GAT GTT AGT TTG GAA ACT TCA Glu Lys Val Ser Lys Ile Ser Pro Cys Asp Val Ser Leu Glu Thr Ser 1940 1945 1950 1955	6093
30	GAT ATA TGT AAA TGT AGT ATA GGG AAG CTT CAT AAG TCA GTC TCA TCT Asp Ile Cys Lys Cys Ser Ile Gly Lys Leu His Lys Ser Val Ser Ser 1960 1965 1970	6141
35	GCA AAT ACT TGT GGG ATT TTT AGC ACA GCA AGT GGA AAA TCT GTC CAG Ala Asn Thr Cys Gly Ile Phe Ser Thr Ala Ser Gly Lys Ser Val Gln 1975 1980 1985	6189
55	GTA TCA GAT GCT TCA TTA CAA AAC GCA AGA CAA GTG TTT TCT GAA ATA Val Ser Asp Ala Ser Leu Gln Asn Ala Arg Gln Val Phe Ser Glu Ile 1990 1995 2000	6237
40	GAA GAT AGT ACC AAG CAA GTC TTT TCC AAA GTA TTG TTT AAA AGT AAC Glu Asp Ser Thr Lys Gln Val Phe Ser Lys Val Leu Phe Lys Ser Asn 2005 2010 2015	6285
45	GAA CAT TCA GAC CAG CTC ACA AGA GAA GAA AAT ACT GCT ATA CGT ACT Glu His Ser Asp Gln Leu Thr Arg Glu Glu Asn Thr Ala Ile Arg Thr 2020 2025 2030 2035	6333
50	CCA GAA CAT TTA ATA TCC CAA AAA GGC TTT TCA TAT AAT GTG GTA AAT Pro Glu His Leu Ile Ser Gln Lys Gly Phe Ser Tyr Asn Val Val Asn 2040 2045 2050	6381
55	TCA TCT GCT TTC TCT GGA TTT AGT ACA GCA AGT GGA AAG CAA GTT TCC Ser Ser Ala Phe Ser Gly Phe Ser Thr Ala Ser Gly Lys Gln Val Ser 2055 2060 2065	6429
33	ATT TTA GAA AGT TCC TTA CAC AAA GTT AAG GGA GTG TTA GAG GAA TTT Ile Leu Glu Ser Ser Leu His Lys Val Lys Gly Val Leu Glu Glu Phe 2070 2075 2080	6477
60	GAT TTA ATC AGA ACT GAG CAT AGT CTT CAC TAT TCA CCT ACG TCT AGA Asp Leu Ile Arg Thr Glu His Ser Leu His Tyr Ser Pro Thr Ser Arg 2085 2090 2095	6525

5					Arg Val		A AAC CCA GAG g Asn Pro Glu 2115	6573
10		Val Asn					A GAA TTT AAA s Glu Phe Lys 2130	6621
10				Val Glu			A AAT AAT CAC 1 Asn Asn His 2145	6669
15	Ser Ile						A GAC AAA CAA n Asp Lys Gln O	6717
20			Gly Thr				C ATT CAT GTT n Ile His Val	6765
25					Lys Asn		G GAA ATT GGT t Glu Ile Gly 2195	6813
		Glu Thr					T ATA GAA GTT n Ile Glu Val 2210	6861
30				Asp Ser			A ACA GAA GCA u Thr Glu Ala 2225	6909
35	Val Glu						G ACA GAT TCT u Thr Asp Ser 0	6957
40		Pro Ser	His Ala				T CCC GAA AAT s Pro Glu Asn	7005
45	GAG GAA Glu Glu 2260	ATG GTT Met Val	TTG TCA Leu Ser 2265	AAT TCA Asn Ser	Arg Ile	GGA AAA AG Gly Lys Ar 270	A AGA GGA GAG g Arg Gly Glu 2275	7053
F.0		. Ile Leu					C TTA TTA AAT n Leu Leu Asn 2290	7101
50				Glu Asn			A AAG GCT TCA u Lys Ala Ser 2305	7149
55							G TTT ATG CAT u Phe Met His 0	7197
60		Ser Leu	Glu Pro				C ACA ACT AAG	7245

5	GAA CGT CAA GAG ATA CAG AAT CCA AAT TTT ACC GCA CCT GGT CAA GAA Glu Arg Gln Glu Ile Gln Asn Pro Asn Phe Thr Ala Pro Gly Gln Glu 2340 2355 2350 2355	7293
J	TTT CTG TCT AAA TCT CAT TTG TAT GAA CAT CTG ACT TTG GAA AAA TCT Phe Leu Ser Lys Ser His Leu Tyr Glu His Leu Thr Leu Glu Lys Ser 2360 2365 2370	7341
10	TCA AGC AAT TTA GCA GTT TCA GGA CAT CCA TTT TAT CAA GTT TCT GCT Ser Ser Asn Leu Ala Val Ser Gly His Pro Phe Tyr Gln Val Ser Ala 2375 2380 2385	7389
15	ACA AGA AAT GAA AAA ATG AGA CAC TTG ATT ACT ACA GGC AGA CCA ACC Thr Arg Asn Glu Lys Met Arg His Leu Ile Thr Thr Gly Arg Pro Thr 2390 2395 2400	7437
20	AAA GTC TTT GTT CCA CCT TTT AAA ACT AAA TCG CAT TTT CAC AGA GTT Lys Val Phe Val Pro Pro Phe Lys Thr Lys Ser His Phe His Arg Val 2405 2410 2415	7485
25	GAA CAG TGT GTT AGG AAT ATT AAC TTG GAG GAA AAC AGA CAA AAG CAA Glu Glu Glu Cys Val Arg Asn Ile Asn Leu Glu Glu Asn Arg Gln Lys Gln 2420 2435 2430 2435	7533
	AAC ATT GAT GGA CAT GGC TCT GAT GAT AGT AAA AAT AAG ATT AAT GAC Asn Ile Asp Gly His Gly Ser Asp Asp Ser Lys Asn Lys Ile Asn Asp 2440 2445 2450	7581
30	AAT GAG ATT CAT CAG TTT AAC AAA AAC AAC TCC AAT CAA GCA GCA GCT Asn Glu Ile His Gln Phe Asn Lys Asn Asn Ser Asn Gln Ala Ala 2455 2460 2465	7629
35	GTA ACT TTC ACA AAG TGT GAA GAA GAA CCT TTA GAT TTA ATT ACA AGT Val Thr Phe Thr Lys Cys Glu Glu Pro Leu Asp Leu Ile Thr Ser 2470 2475 2480	7677
40	CTT CAG AAT GCC AGA GAT ATA CAG GAT ATG CGA ATT AAG AAG AAA CAA Leu Gln Asn Ala Arg Asp Ile Gln Asp Met Arg Ile Lys Lys Gln 2485 2490 2495	7725
45	AGG CAA CGC GTC TTT CCA CAG CCA GGC AGT CTG TAT CTT GCA AAA ACA Arg Gln Arg Val Phe Pro Gln Pro Gly Ser Leu Tyr Leu Ala Lys Thr 2500 2505 2510 2515	7773
	TCC ACT CTG CCT CGA ATC TCT CTG AAA GCA GCA GTA GGA GGC CAA GTT Ser Thr Leu Pro Arg Ile Ser Leu Lys Ala Ala Val Gly Gly Gln Val 2520 2525 2530	7821
50	CCC TCT GCG TGT TCT CAT AAA CAG CTG TAT ACG TAT GGC GTT TCT AAA Pro Ser Ala Cys Ser His Lys Gln Leu Tyr Thr Tyr Gly Val Ser Lys 2535 2540 2545	7869
55	CAT TGC ATA AAA ATT AAC AGC AAA AAT GCA GAG TCT TTT CAG TTT CAC His Cys Ile Lys Ile Asn Ser Lys Asn Ala Glu Ser Phe Gln Phe His 2550 2555 2560	7917
60	ACT GAA GAT TAT TTT GGT AAG GAA AGT TTA TGG ACT GGA AAA GGA ATA Thr Glu Asp Tyr Phe Gly Lys Glu Ser Leu Trp Thr Gly Lys Gly Ile 2565 2570 2575	7965
	CAG TTG GCT GAT GGT GGA TGG CTC ATA CCC TCC AAT GAT GGA AAG GCT	8013

	Gln 2580	Leu	Ala	Asp	_	Gly 2585	Trp	Leu	Ile		Ser 2590	Asn	Asp	Gly	Lys 2	Ala :595	
5				Glu					Leu					Gly	GTG Val 2610		8061
10			Leu					Trp					Tyr		TGG Trp		8109
15		Trp					Met					Pro			TTT Phe		8157
20	Asn					Pro					Leu				TAC Tyr		8205
					Ile					Arg					AAG Lys		8253
25				Asp					Lys					Cys	GTT Val 2690		8301
30			Ile					Asn					Ser		AAT Asn		8349
35		Ser					Gln					Ile			ACA Thr		8397
40	Gly					Lys					Pro				GCT Ala		8445
		Lys			Arg					Gln					CAT His		8493
45				Val					Ala					Glu	GCC Ala 2770		8541
50			Leu					Ser					Arg		GCT Ala		8589
55		Tyr					Phe					Arg			CCT Pro		8637
60	Pro					Phe					Asn				GTT Val		8685
~ <b>~</b>															ACA Thr		8733

3060

3070

Phe Leu Asp Pro Asp Phe Gln Pro Ser Cys Ser Glu Val Asp Leu Ile

3065

5				GTT Val					Lys					Pro		_	9501
10			Ser	GAC Asp 3095				Asn					Lys				9549
10		Leu		GAG Glu			Ile					Leu					9597
15	Asn			TGG Trp		Pro					Gly						9645
20				TTT Phe	Ser					Ser					His		9693
25				TTC Phe					Asn					Ile			9741
30			Asn	GAA Glu 3175				Lys					Leu				9789
30		Pro		TGG Trp			Pro					Thr					9837
35	Thr			ATC Ile		Pro					Lys						9885
40				GAG Glu	Ile					Pro					Met		9933
45				TCT Ser					Val					Thr			9981
50			Lys	GGG Gly 3255				Ile					Asn				10029
30		Arg		TTG Leu	_	_	Leu					Leu					10077
55	Ser			TGT Cys		Phe					Ala						10125
60		Pro		AGT Ser	Cys					Glu					Lys		10173

5				TCT Ser					Pro					Asn			10221
_			Leu	GAA Glu 3335				Ile					Leu				10269
10		Thr		GCT Ala			Ser					Glu					10317
15	Ser			GAA Glu		Thr					Thr						10365
20				AAA Lys	Arg					Ser					Gln		10413
25				GCC Ala					Cys					Gln			10461
			Thr	AAA Lys 3415				TAA									10485
30			(2)	) INI	FORMA	IOITA	N FO	R SE(	Q ID	NO:	11:						
35		( :	(A) (B) (C)	EQUEN LENC TYPI STRA TOPC	ETH: E: ar ANDEI	3418 mino ONESS	3 am: acio 3: s:	ino a i ingle	acid	70							
35		(:	(A) (B) (C) (D)	LENG TYPE STRA TOPG MOLEG	ETH: E: ar ANDEI OLOGY CULE ENT	3418 mino ONESS Y: 1: TYPE	acions: sinear	ino a ingle cote: cerna	acids e in al		NO.	11.					
	Met	(:	(A) (B) (C) (D) ii) N v) FN	LENC TYPE STRA TOPO MOLEC RAGME	ETH: E: ar ANDEI DLOGY CULE ENT T	3418 nino DNESS Y: l: TYPE TYPE DESC	acics: s: s: inear	ino a ingle rote: cerna	e in al : SE	QI Ç			Glu	Īļe	Dhe	Iws	
	1	(; (v ()	(A) (B) (C) (D) ii) N v) FI xi) S	LENG TYPE STRA TOPO MOLEC RAGME SEQUE	ETH: E: ar ANDEL DLOGY CULE ENT T ENCE Ser 5	3418 mino DNESS Y: 1: TYPE TYPE DESC	acic acic S: s: inear c: pr : int CRIPT	ino a ingle cote: cote: cote: TION	in al Pro	Q ID Thr 10	Phe	Phe			15	•	
40	1 Thr	(; (; Pro Arg	(A) (B) (C) (D) iii) f (V) FF (Xi) S Ile Cys	LENC TYPE STRA TOPO MOLEC RAGME	ETH: E: ar ANDEI DLOGY CULE ENT T ENCE Ser 5	3418 mino DNESS Y: 1: TYPE TYPE DESC Lys Ala	acic acic S: s: inear E: pr : int CRIPT	ino a ingle ingle cote cote corn TION Arg Leu	in al SEG	Q ID Thr 10 Pro	Phe Ile	Phe Ser	Leu	Asn 30	15 Trp	Phe	
40	l Thr Glu	(; (x Pro Arg	(A) (B) (C) (D)  iii) P (Xi) S  Ile  Cys  Leu 35	LENG TYPE STRA TOPO MOLEG RAGME SEQUE Gly Asn 20	GTH: E: ar ANDER DLOGY CULE ENT T ENCE Ser Lys Ser	3418 mino DNESS Y: 1: TYPE TYPE DESC Lys Ala Glu	acic acic s: s: inear crip crip Glu Asp Ala	ino a ingle ingle cote: cote: corn ingle cote: c	in al Pro Gly 25 Pro	Q ID Thr 10 Pro	Phe Ile Asn	Phe Ser Ser	Leu Glu 45	Asn 30 Pro	15 Trp Ala	Phe Glu	
40 45	1 Thr Glu Glu Pro	(: (x (x Pro Arg Glu Ser 50	(A) (B) (C) (D)  ii) F  xi) S  Ile  Cys  Leu 35 Glu	LENG TYPH STRA TOPO MOLEO RAGMH SEQUH Gly Asn 20 Ser	ETH: E: ar ANDEI OLOGY CULE ENT T ENCE Ser 5 Lys Ser Lys	3418 mino DNESS Y: 1: TYPE DESC Lys Ala Glu Asn Ser	acic scinear cripe cripe Glu Asp Ala Asn 55	ino a ingle cote: cote: cote: corn TION Arg Leu Pro 40 Asn	in al Pro Gly 25 Pro Tyr	Q ID Thr 10 Pro Tyr Glu	Phe Ile Asn Pro	Phe Ser Ser Asn 60	Leu Glu 45 Leu	Asn 30 Pro	15 Trp Ala Lys	Phe Glu Thr	
40 45 50	1 Thr Glu Glu Pro 65	(: (x) (x) Pro Arg Glu Ser 50 Gln	(A) (B) (C) (C) (D)  ii) F  xi) S  Ile Cys Leu 35 Glu Arg	LENG TYPH STRA TOPO MOLEC RAGMH SEQUH Gly Asn 20 Ser His	ETH: E: ar ANDEI DLOGY CULE ENT T ENCE Ser 5 Lys Ser Lys Pro	3418 mino DNESS Y: 1: TYPE TYPE DESC Lys Ala Glu Asn Ser 70	acic acic s: s: inear crip: Glu Asp Ala Asn 55 Tyr	ino a ingle ingle cote:	in al SEG Pro Gly 25 Pro Tyr Gln	Thr 10 Pro Tyr Glu Leu Leu	Phe Ile Asn Pro Ala 75	Phe Ser Ser Asn 60 Ser	Leu Glu 45 Leu Thr	Asn 30 Pro Phe Pro	15 Trp Ala Lys Ile Val	Phe Glu Thr Ile	
40 45	1 Thr Glu Glu Pro 65 Phe	(: (x) (x) Pro Arg Glu Ser 50 Gln Lys	(A) (B) (C) (C) (D)  ii) Fi  xi) S  Ile Cys Leu 35 Glu Arg Glu	LENG TYPE STRA TOPO MOLEG RAGME SEQUE Gly Asn 20 Ser His Lys Gln Lys	ETH: E: ar ANDER OLOGY CULE ENT : ENCE Ser 5 Lys Ser Lys Pro Gly 85	3418 mino DNESS Y: 1: TYPE TYPE DESC Lys Ala Glu Asn Ser 70 Leu	acic acic s: s: inear crip: crip: Glu Asp Ala Asn 55 Tyr Thr	ino a dingle ingle cote:	in al Ero Gly 25 Pro Tyr Gln Pro Leu	Q ID Thr 10 Pro Tyr Glu Leu Leu 90	Phe Ile Asn Pro Ala 75 Tyr	Phe Ser Ser Asn 60 Ser Gln	Leu Glu 45 Leu Thr	Asn 30 Pro Phe Pro Pro	15 Trp Ala Lys Ile Val 95	Phe Glu Thr Ile 80 Lys	
40 45 50	1 Thr Glu Glu Pro 65 Phe	(: (x) (x) Pro Arg Glu Ser 50 Gln Lys Leu	(A) (B) (C) (C) (D)  ii) fr  xi) S  Ile Cys Leu 35 Glu Arg Glu Asp Lys	LENG TYPH STRA TOPO MOLEG RAGMH SEQUH Gly Asn 20 Ser His Lys	ETH: E: ar ANDEI DLOGY CULE ENT T ENCE Ser Lys Ser Lys Pro Gly 85 Phe	3418 mino DNESS Y: 1: TYPE DESC Lys Ala Glu Asn Ser 70 Leu Lys	acic acic S: s: inear E: pr : int CRIPT Glu Asp Ala Asn 55 Tyr Thr	ino a dingle ingle cotes cerna fion Arg Leu Pro Asn Asn Leu Asp Val	in al Ero Gly 25 Pro Tyr Gln Pro Leu 105	Thr 10 Pro Tyr Glu Leu 90 Gly	Phe Ile Asn Pro Ala 75 Tyr Arg	Phe Ser Ser Asn 60 Ser Gln Asn	Leu Glu 45 Leu Thr Ser Val Asp	Asn 30 Pro Phe Pro Pro	15 Trp Ala Lys Ile Val 95 Asn	Phe Glu Thr Ile 80 Lys Ser	
40 45 50	1 Thr Glu Glu Pro 65 Phe Glu Arg	(; (x) (x) Pro Arg Glu Ser 50 Gln Lys Leu His	(A) (B) (C) (C) (D)  ii) Fr  xi) S  Ile Cys Leu 35 Glu Arg Glu Asp Lys 115	LENGTYPE STRATOPO MOLEG RAGME SEQUE Gly Asn 20 Ser His Lys Gln Lys 100	ETH: E: ar ANDEI DLOGY CULE ENT T ENCE Ser Lys Ser Lys Pro Gly 85 Phe Leu	3418 mino DNESS Y: 1: TYPE DESC Lys Ala Glu Asn Ser 70 Leu Lys Arg	acic acic S: s: inear E: pr : int CRIPT Glu Asp Ala Asn 55 Tyr Thr Leu	ino a dingle ingle cotes cerna fion Arg Leu Pro Asn Asn Leu Asp Val 120	in al SEG Pro Gly 25 Pro Tyr Gln Pro Leu 105 Lys	Thr 10 Pro Tyr Glu Leu 90 Gly	Phe Ile Asn Pro Ala 75 Tyr Arg Lys	Phe Ser Ser Asn 60 Ser Gln Asn Met	Leu Glu 45 Leu Thr Ser Val Asp 125	Asn 30 Pro Phe Pro Pro 110 Gln	15 Trp Ala Lys Ile Val 95 Asn Ala	Phe Glu Thr Ile 80 Lys Ser Asp	

						150					155					1.00
	145	Clar	Sar	Len	Dhe	150	Thr	Pro	Larg	Dhe	155 Val	Laze	Glv	Arg	Gln	160 Thr
<b>r</b> -	_	-			165					170					175	
5				180					185					Pro 190		
		-	195					200					205	Ser		
10	Leu	Ile 210	Val	Arg	Asn	Glu	Glu 215	Ala	Ser	Glu	Thr	Val 220	Phe	Pro	His	Asp
	Thr 225	Thr	Ala	Asn	Val	Lys 230	Ser	Tyr	Phe	Ser	Asn 235	His	Asp	Glu	Ser	Leu 240
		Lys	Asn	Asp	Arg 245	Phe	Ile	Ala	Ser	Val 250	Thr	Asp	Ser	Glu	Asn 255	Thr
15	Asn	Gln	Arg	Glu 260	Ala	Ala	Ser	His	Gly 265	Phe	Gly	Lys	Thr	Ser 270	Gly	Asn
	Ser	Phe	Lys 275	Val	Asn	Ser	Cys	Lys 280	Asp	His	Ile	Gly	Lys 285	Ser	Met	Pro
20	Asn	Val 290	Leu	Glu	Asp	Glu	Val 295	Tyr	Glu	Thr	Val	Val 300	Asp	Thr	Ser	Glu
	Glu 305	Asp	Ser	Phe	Ser	Leu 310	Cys	Phe	Ser	Lys	Cys 315	Arg	Thr	Lys	Asn	Leu 320
	Gln	Lys	Val	Arg	Thr 325	Ser	Lys	Thr	Arg	Lys 330	Lys	Ile	Phe	His	Glu 335	Ala
25	Asn	Ala	Asp	Glu 340	Cys	Glu	Lys	Ser	Lys 345	Asn	Gln	Val	Lys	Glu 350	Lys	Tyr
	Ser	Phe	Val 355	Ser	Glu	Val	Glu	Pro 360	Asn	Asp	Thr	Asp	Pro 365	Leu	Asp	Ser
30	Asn	Val 370	Ala	His	Gln	Lys	Pro 375	Phe	Glu	Ser	Gly	Ser 380	Asp	Lys	Ile	Ser
	Lys 385	Glu	Val	Val	Pro	Ser 390	Leu	Ala	Cys	Glu	Trp 395	Ser	Gln	Leu	Thr	Leu 400
	Ser	Gly	Leu	Asn	Gly 405	Ala	Gln	Met	Glu	Lys 410	Ile	Pro	Leu	Leu	His 415	Ile
35				420					425	_				Asp 430		
			435		_	_		440					445	Leu		
40		450					455					460		Glu		
	Val 465	Asn	Lys	Arg	Asp	Glu 470	Glu	Gln	His	Leu	Glu 475	Ser	His	Thr	Asp	Cys 480
	Ile	Leu	Ala	Val	Lys 485	Gln	Ala	Ile	Ser	Gly 490	Thr	Ser	Pro	Val	Ala 495	Ser
45				500					505					Glu 510		
	Lys	Glu	Thr 515	Phe	Asn	Ala	Ser	Phe 520	Ser	Gly	His	Met	Thr 525	Asp	Pro	Asn
50	Phe	Lys 530	Lys	Glu	Thr	Glu	Ala 535	Ser	Glu	Ser	Gly	Leu 540	Glu	Ile	His	Thr
	Val 545	Cys	Ser	Gln	Lys	Glu 550	Asp	Ser	Leu	Cys	Pro 555	Asn	Leu	Ile	Asp	Asn 560
		Ser	Trp	Pro	Ala 565		Thr	Thr	Gln	Asn 570		Val	Ala	Leu	Lys 575	
55	Ala	Gly	Leu	Ile 580	Ser	Thr	Leu	Lys	Lys 585		Thr	Asn	Lys	Phe 590	Ile	Tyr
	Ala	Ile	His 595	Asp	Glu	Thr	Ser	Tyr 600	Lys	Gly	Lys	Lys	Ile 605	Pro	Lys	Asp
60	Gln	Lys 610	Ser	Glu	Leu	Ile	Asn 615	Cys	Ser	Ala	Gln	Phe 620	Glu	Ala	Asn	Ala
	Phe 625	Glu	Ala	Pro	Leu	Thr 630	Phe	Ala	Asn	Ala	Asp 635	Ser	Gly	Leu	Leu	His 640

	Ser	Ser	Val	Lys	Arg 645	Ser	Cys	Ser	Gln	Asn 650	Asp	Ser	Glu	Glu	Pro 655	Thr
5	Leu	Ser	Leu	Thr 660		Ser	Phe	Gly	Thr 665		Leu	Arg	Lys	Cys 670		Arg
J	Asn	Glu	Thr 675		Ser	Asn	Asn	Thr 680		Ile	Ser	Gln	Asp		Asp	Tyr
	Lys	Glu 690		Lys	Cys	Asn	Lys 695		Lys	Leu	Gln	Leu 700		Ile	Thr	Pro
10	Glu 705	Ala	Asp	Ser	Leu	Ser 710	Cys	Leu	Gln	Glu	Gly 715	Gln	Cys	Glu	Asn	Asp 720
	Pro	Lys	Ser	Lys	Lys 725	Val	Ser	Asp	Ile	Lys 730	Glu	Glu	Val	Leu	Ala 735	Ala
15	Ala	Cys	His	Pro 740	Val	Gln	His	Ser	Lys 745	Val	Glu	Tyr	Ser	Asp 750	Thr	Asp
	Phe	Gln	Ser 755	Gln	Lys	Ser	Leu	Leu 760	Tyr	Asp	His	Glu	Asn 765	Ala	Ser	Thr
		Ile 770					775					780				
20	785	Ser	_	_	-	790		_	_		795				_	800
	Asn	Asn	Tyr	Glu	Ser 805	Asp	Val	Glu	Leu	Thr 810	Lys	Asn	Ile	Pro	Met 815	Glu
25	_	Asn		820		_			825			_	_	830		
		Leu	835			-	-	840	_				845		_	-
		Gln 850					855					860				
30	865	Glu				870		_			875			_		880
		Leu			885					890					895	
35		Arg		900					905		_			910		
	_	Leu -	915	_				920			_		925			
4.0		Tyr 930		_			935	_				940				_
40	945	Asp			_	950					955	_				960
		His			965			_		970				_	975	
45		Asn		980	-				985			-	-	990		-
		Ala	995					1000	)				1005	5	_	
50		Arg 1010 Lys	)				1015	5		_		1020	)			
50	102	5				1030	)				1035	5		_		104
		Leu		_	1045	5				1050	)			-	1055	5
55		Lys		1060	)				1065	5				1070	)	
		Ser	1075	5				1080	) -	-			1085	5		
60		Met 1090	)				1095	5				1100	)			
00	110					1110	כ				1115	5				112
	ser	Gly	ser	GIN	rne	GIU	rne	Thr	GIN	rne	arg	ьys	Pro	ser	Tyr	тте

					1125					1130					1135	
	Leu	Gln	Lys	Ser 1140		Phe	Glu	Val	Pro 1145		Asn	Gln	Met	Thr 1150		Leu
5	Lys	Thr	Thr 1155	Ser	Glu	Glu	Cys	Arg 1160		Ala	Asp	Leu	His 1165		Ile	Met
	Asn	Ala 1170		Ser	Ile	Gly	Gln 1175		Asp	Ser	Ser	Lys 1180		Phe	Glu	Gly
10	Thr 1185		Glu	Ile	_	Arg 1190	_	Phe	Ala	Gly	Leu 1195		Lys	Asn	Asp	Cys 120
			Ser	Ala		Gly		Leu		Asp 1210		Asn	Glu	Val	Gly 1215	
	Arg	Gly	Phe	Tyr 1220	Ser		His	Gly	Thr 1225		Leu	Asn	Val	Ser 1230		Glu
15	Ala	Leu	Gln 1235	Lys	Ala	Val	Lys	Leu 1240		Ser	Asp	Ile	Glu 1245		Ile	Ser
	Glu	Glu 1250		Ser	Ala		Val 1255		Pro	Ile	Ser	Leu 1260		Ser	Ser	Lys
20	Cys 1265		Asp	Ser		Val 1270		Met	Phe	Lys	Ile 1275		Asn	His	Asn	Asp 128
	Lys	Thr	Val	Ser	Glu 1285		Asn	Asn	Lys	Cys 1290		Leu	Ile	Leu	Gln 1295	
	Asn	Ile	Glu	Met 1300		Thr	Gly	Thr	Phe 1305	Val		Glu	Ile	Thr 1310		Asn
25	Tyr	Lys	Arg 1315	Asn	Thr	Glu	Asn	Glu 1320		Asn	Lys	Tyr	Thr 1325		Ala	Ser
	Arg	Asn 1330		His	Asn	Leu	Glu 1339		Asp	Gly	Ser	Asp 1340		Ser	Lys	Asn
30	Asp 1345	Thr		Cys	Ile	His 1350		Asp	Glu	Thr	Asp 1355		Leu	Phe	Thr	Asp 136
			Asn	Ile	Cys 1365		Lys	Leu	Ser	Gly 1370		Phe	Met	Lys	Glu 1375	
	Asn	Thr	Gln	Ile 1380		Glu	Asp	Leu	Ser 1385		Leu	Thr	Phe	Leu 1390		Val
35	Ala	Lys	Ala 1399	Gln 5	Glu	Ala	Cys	His 1400		Asn	Thr	Ser	Asn 140		Glu	Gln
	Leu	Thr 1410		Thr	Lys	Thr	Glu 141		Asn	Ile	Lys	Asp 142		Glu	Thr	Ser
40	1425	5		Phe		1430	)				143	5				144
				Asn	1445	5				1450	0				1455	5
				Phe 1460	)				1465	5				1470	0	
45			147	_				1480	С				148	5		
	-	1490	С	Lys			149	5				150	0			
50	1509	5		Gly		1510	)				151	5				152
			_	Phe	1525	5				153	0				153	5
				Asp 1540	)				154	5				155	0	
55			155					1560	0		-		156	5		
		157	0	Ala			157	5				158	0			
60	Ile 158		Ala	Ala	Pro	Lys 1590		Lys	Glu	Met	Gln 159		Ser	Leu	Asn	Asn 160
	Asp	Lys	Asn	Leu	Val 160		Ile	Glu	Thr	Val 161		Pro	Pro	Lys	Leu 161	

	Ser	Asp	Asn	Leu 1620		Arg	Gln	Thr	Glu 1625		Leu	Lys	Thr	Ser 1630		Ser
5	Ile	Phe	Leu 1635		Val	Lys		His 1640		Asn	Val	Glu	Lys 1645		Thr	Ala
	Lys	Ser 1650		Ala	Thr	Cys	Tyr 1655		Asn	Gln	Ser	Pro 1660		Ser	Val	Ile
	Glu 1669	Asn	Ser	Ala	Leu	Ala 1670		Tyr	Thr	Ser	Cys 1675		Arg	Lys	Thr	Ser 168
10		Ser	Gln	Thr	Ser 1685		Leu	Glu	Ala	Lys 1690		Trp	Leu	Arg	Glu 1695	
	Ile	Phe	Asp	Gly 1700		Pro	Glu	Arg	Ile 1709		Thr	Ala	Asp	Tyr 1710		Gly
15		Tyr	1715	5				1720	)				1725	5		
	Lys	Asn 1730		Leu	Ser	Glu	Lys 1735		Asp	Thr	Tyr	Leu 1740		Asn	Ser	Ser
	Met 174!	Ser 5	Asn	Ser	Tyr	Ser 1750		His	Ser	Asp	Glu 1755		Tyr	Asn	Asp	Ser 176
20	-	Tyr			1765	5	_			1770	)				1775	5
	_	Asn		1780	)				1785	5				1790	)	
25		Val	1795	5				1800	)				1809	5		
	-	Val 1810	)				1819	5				1820	0			
	182			-		1830	)				1835	5				184
30		Pro			1845	5			_	1850	)				1855	5
		Thr		1860	)		_		186	5				1870	)	
35		Ile	1879	5				1880	)				188	5		
		Met 1890	)	_	_		189	5				190	0			
4.0	190					191	С				191	5				192
40		Ala	_		1925	5				1930	)				193	5
		Gly		1940	o -			-	194	5				1950	)	
45		Thr	195	5		_	_	1960	0				196	5		
		Ser 1970	)				197	5				198	0			
50	198	Val 5 Glu				199	0				199	5				200
30		Ser			200	5				2010	С				201	5
	_	Arg		2020	0				202	5				203	0	
55		Val	203	5				204	0				204	5		
		205 Val	0				205	5	_			206	0			
60	206					207	0				207	5				208
		Ser		_	208	5				209	0				209	5
			_					-				_		_	-	_

		2100		2105	2110	
	211	5	2120	)	Lys Thr Cys S 2125	_
5	2130		2135		Gly Gly Ser S 2140	
	2145	215	0	215		216
10		2165		2170		2175
		2180		2185	Lys Asn Val 1 2190	
7.5	219	5	220	O .	Pro Val Lys 2205	
15	2210		2215		Glu Asn Tyr I	
	2225	223	10	223		224
20	-	2245		2250		2255
		2260		2265	Arg Ile Gly 1 2270	
٥٦	227	5	228	0	Ser Ile Lys 2 2285	
25	2290		2295		Gln Glu Lys 3	
	2305	231	LO	231		232
30		2325		2330		2335
	_	2340		2345	Asn Phe Thr 2350	
2.5	235	5	236	0	Glu His Leu ' 2365	
35	2370		2375		His Pro Phe '2380	
	2385	239	90	239		240
40	•	2405		2410		2415
	_	2420		2425	Leu Glu Glu 2430	
4.5	243	5	244	0	Asp Ser Lys . 2445	
45	2450		2455		Asn Asn Ser . 2460	
	2465	24	70	247		248
50		2485		2490		2495
		2500		2505	Gly Ser Leu 2510	
	251	.5	252	0	Lys Ala Ala	
55	2530		2535		Leu Tyr Thr 2540	
	2545	25	50	255		256
60		2565		2570		2575
	гàs Già lie	e Gin Leu Ala 2580	a Asp Gly	Gly Trp Leu 2585	Ile Pro Ser 2590	

	Gly	Lys	Ala 2599	-	Lys	Glu	Glu	Phe 2600		Arg	Ala	Leu	Cys 2605	-	Thr	Pro
5	Gly	Val 261		Pro	Lys	Leu	Ile 261	Ser		Ile	Trp	Val 2620		Asn	His	Tyr
	Arg 2629		Ile	Ile	Trp	Lys 2630		Ala	Ala	Met	Glu 2639		Ala	Phe	Pro	Lys 264
	Glu	Phe	Ala	Asn	Arg 2645		Leu	Ser	Pro	Glu 2650		Val	Leu	Leu	Gln 2655	
10	Lys	Tyr	Arg	Tyr 2660	Asp	Thr	Glu	Ile	Asp 266	_	Ser	Arg	Arg	Ser 2670		Ile
			2675	5	Glu			2680	C				2685	5		
15		2690	)		Ile		2695	5				2700	)			
	2705	5			Ser	2710	)				2715	5				272
0.0					Trp 2725	5				2730	)				2735	5
20				2740					2745	5		_		2750	)	
			2755	5	Glu			2760	)				2765	5		
25		2770	)		Ser Tyr		2775	5	_			2780	)			_
	2785	5			Leu	2790	)				2795	5				280
30					2805 Ile	5				2810	)		_		2815	5
50				2820					2825	5				2830	)	
			2835	5	Lys			2840	)				2845	5		
35		2850	)		Ile		2855	5				2860	)			
	2865	5			Leu	2870	)				2875	5				288
40					2885 Gly	5				2890	)				2895	5
	Asp	Pro	Ala	2900 Tyr	Leu	Glu	Gly	Tyr	2909 Phe		Glu	Glu	Gln	2910 Leu		Ala
	Leu	Asn	2915 Asn		Arg		Met			Asp					Gln	Ile
45				Ile	Arg				Glu	Ser				Lys	Glu	Gln
	2945 Gly		Ser	Arg	Asp			Thr	Val				Arg	Ile	Val	296 Ser
50	Tyr	Ser	Lys		2965 Glu		Asp	Ser				Ser	Ile		-	
	Ser	Ser	Asp 2995		) Tyr	Ser	Leu				Gly	Lys				Ile
55	Tyr	His 3010	Leu		Thr	Ser	Lys 3015			Ser	Lys	Ser 3020			Ala	Asn
	Ile 3025	Gln		Ala	Ala	Thr 3030	Lys		Thr	Gln	Tyr 3035	Gln		Leu	Pro	Val 304
			Glu	Ile	Leu 3045	Phe		Ile	Tyr	Gln 3050	Pro		Glu	Pro	Leu 3055	His
60	Phe	Ser	Lys	Phe 3060	Leu		Pro	Asp	Phe 3065	Gln		Ser	Cys	Ser 3070	Glu	
	Acr	Lou	TIG		Dho	1751	7727	000			T ***	T	mb ac			77-

			30/3	)				3080	,				3083	•		
	Pro	Phe 3090		Tyr	Leu	Ser	Asp 3099		Cys	Tyr	Asn	Leu 3100		Ala	Ile	Lys
5	Phe 3105	Trp		Asp	Leu	Asn 3110	Glu		Ile	Ile	Lys 3115	Pro		Met	Leu	Ile 312
		Ala	Ser	Asn	Leu			Ara	Pro	Glu			Ser	Glv	Leu	
					3125		1-	5		3130		-1-		1	3135	
10	Thr	Leu	Phe	Ala 3140	_	Asp	Phe	Ser	Val 3145		Ser	Ala	Ser	Pro 3150	_	Glu
	Gly	His	Phe 3155		Glu	Thr	Phe	Asn 3160		Met	Lys	Asn	Thr 3169		Glu	Asn
	Ile	Asp 3170		Leu	Cys	Asn	Glu 3179		Glu	Asn	Lys	Leu 3180		His	Ile	Leu
15	His 3189	Ala		Asp	Pro	Lys 3190	Trp		Thr	Pro	Thr 3195	Lys		Cys	Thr	Ser
		Pro	Tyr	Thr	Ala			Ile	Pro	Gly			Asn	Lys	Leu	
			_		3205	5				3210	)				3219	5
20	Met	Ser	Ser	Pro 3220		Cys	Glu	Ile	Tyr 3225		Gln	Ser	Pro	Leu 3230		Leu
	Cys	Met	Ala 3235	_	Arg	Lys	Ser	Val 3240		Thr	Pro	Val	Ser 3245		Gln	Met
	Thr	Ser 3250		Ser	Cys	Lys	Gly 3259		Lys	Glu	Ile	Asp 3260		Gln	Lys	Asn
25	Cys 3269	Lys	Lys	Arg	Arg	Ala 3270	Leu		Phe	Leu	Ser 3275	Arg		Pro	Leu	Pro
		Pro	Val	Ser	Pro 3285	Ile		Thr	Phe	Val 3290	Ser		Ala	Ala	Gln 3295	Lys
	Ala	Phe	Gln	Pro			Ser	Cvs	Glv			Tvr	Glu	Thr		
30				3300	)				3305	5	_	_		3310	)	
	Lys	Lys	Lys 3315		Leu	Asn	Ser	Pro 3320		Met	Thr	Pro	Phe 3325	-	Lys	Phe
		Glu 3330	)				3335	5				3340	)			
35	Ala 3345	Leu 5	Ile	Asn	Thr	Gln 3350		Leu	Leu	Ser	Gly 3355		Thr	Gly	Glu	Lys 336
	Gln	Phe	Ile	Ser			Glu	Ser	Thr			Ala	Pro	Thr		
	Glu	Asp	Tvr	Leu	3369 Ara		Lvs	Ara	Ara	3370 Cvs		Thr	Ser	Leu	3375	
40				3380	)				3385	5				3390	)	
	Glu	Gln	Glu 3399		Ser	Gln	Ala	Ser 3400		Glu	Glu	Cys	Glu 3405		Asn	Lys
	Gln	Asp 3410		Ile	Thr	Thr	Lys 3419	_	Tyr	Ile						
45			(0)		IODM7											
			(2)	INF	ORMA	4.1.1.OI	V FOI	K SEÇ	מד ו	NO:	12:					
		( j		EQUEN LENC												
50			(B)	TYPE	E: ni	ıclei	ic ac	cid		•						
				STRA					<u> </u>							
			(1)	1010	LOGI	(; <u>1</u> ]	Liieai	-								
55				OLEC FEATU		TYPE	E: cI	ONA								
J J		( )														
				NAN LOC						ice						
				OTE						A2 (C	MI5)					
60		1-														
		( )	(T)	SEQUE	NUL	しほど	YXTD.	L TON:	SEC	תד ג	MO: ]	∠:				

5	TCTGC	CTGC ATTI	GC C	TCGC	GTG1	C TI	TTTGC	CGGCC CGTCA	GTC GCT	GGT(	GCC CCCG	GCCI GCCI	GGAC AAAA	SAA ( AAG A A AT(	CGTC ACTC CCI	GGCGCC BAGGGG BCACCT CATT O Ile	60 120 180 237
10	GGA T																285
15	AAC A Asn I 20																333
20	TCT T																381
	CAT A																429
25	AAA (																477
30	CAA ( Gln (																525
35	AAA 1 Lys I 100																573
40	AGT (																621
10	TGT (																669
45	TGT A																717
50	TTG T																765
55	ATT TILE S																813
60	AGT :																861
	AGA A																909

AAT GTG AAA AGC TAT TTT TCC AAT CAT GAT GAA AGT CTG AAG AAA AAT Asn Val Lys Ser Tyr Phe Ser Asn His Asp Glu Ser Leu Lys Lys Asn GAT AGA TTT ATC GCT TCT GTG ACA GAC AGT GAA AAC ACA AAT CAA AGA Asp Arq Phe Ile Ala Ser Val Thr Asp Ser Glu Asn Thr Asn Gln Arq GAA GCT GCA AGT CAT GGA TTT GGA AAA ACA TCA GGG AAT TCA TTT AAA Glu Ala Ala Ser His Gly Phe Gly Lys Thr Ser Gly Asn Ser Phe Lys GTA AAT AGC TGC AAA GAC CAC ATT GGA AAG TCA ATG CCA CAT GTC CTA Val Asn Ser Cys Lys Asp His Ile Gly Lys Ser Met Pro His Val Leu GAA GAT GAA GTA TAT GAA ACA GTT GTA GAT ACC TCT GAA GAA GAT AGT Glu Asp Glu Val Tyr Glu Thr Val Val Asp Thr Ser Glu Glu Asp Ser TTT TCA TTA TGT TTT TCT AAA TGT AGA ACA AAA AAT CTA CAA AAA GTA Phe Ser Leu Cys Phe Ser Lys Cys Arg Thr Lys Asn Leu Gln Lys Val AGA ACT AGC AAG ACT AGG AAA AAA ATT TTC CAT GAA GCA AAC GCT GAT Arg Thr Ser Lys Thr Arg Lys Lys Ile Phe His Glu Ala Asn Ala Asp GAA TGT GAA AAA TCT AAA AAC CAA GTG AAA GAA AAA TAC TCA TTT GTA Glu Cys Glu Lys Ser Lys Asn Gln Val Lys Glu Lys Tyr Ser Phe Val TCT GAA GTG GAA CCA AAT GAT ACT GAT CCA TTA GAT TCA AAT GTA GCA Ser Glu Val Glu Pro Asn Asp Thr Asp Pro Leu Asp Ser Asn Val Ala CAT CAG AAG CCC TTT GAG AGT GGA AGT GAC AAA ATC TCC AAG GAA GTT His Gln Lys Pro Phe Glu Ser Gly Ser Asp Lys Ile Ser Lys Glu Val GTA CCG TCT TTG GCC TGT GAA TGG TCT CAA CTA ACC CTT TCA GGT CTA Val Pro Ser Leu Ala Cys Glu Trp Ser Gln Leu Thr Leu Ser Gly Leu AAT GGA GCC CAG ATG GAG AAA ATA CCC CTA TTG CAT ATT TCT TCA TGT Asn Gly Ala Gln Met Glu Lys Ile Pro Leu Leu His Ile Ser Ser Cys GAC CAA AAT ATT TCA GAA AAA GAC CTA TTA GAC ACA GAG AAC AAA AGA Asp Gln Asn Ile Ser Glu Lys Asp Leu Leu Asp Thr Glu Asn Lys Arg AAG AAA GAT TTT CTT ACT TCA GAG AAT TCT TTG CCA CGT ATT TCT AGC Lys Lys Asp Phe Leu Thr Ser Glu Asn Ser Leu Pro Arg Ile Ser Ser CTA CCA AAA TCG GAG AAG CCA TTA AAT GAG GAA ACA GTG GTA AAT AAG 

Leu Pro Lys Ser Glu Lys Pro Leu Asn Glu Glu Thr Val Val Asn Lys

5	GAT Asp								1677
10	AAG Lys 485								1725
10	ATC Ile								1773
15	AAT Asn								1821
20	ACT Thr								1869
25	AAG Lys								1917
30	GCC Ala 565								1965
30	TCC Ser								2013
35	GAA Glu								2061
40	CTA Leu								2109
45	CTT Leu								2157
50	AGA Arg 645								2205
20	AGC Ser								2253
55	TCT Ser								2301
60	TGT Cys								2349

_			TCA Ser 710														2397
5	AAA Lys	AAA Lys 725	GTT Val	TCA Ser	GAT Asp	ATA Ile	AAA Lys 730	GAA Glu	GAG Glu	GTC Val	TTG Leu	GCT Ala 735	GCA Ala	GCA Ala	TGT Cys	CAC His	2445
10	CCA Pro 740	GTA Val	CAA Gln	CAC His	TCA Ser	AAA Lys 745	GTG Val	GAA Glu	TAC Tyr	AGT Ser	GAT Asp 750	ACT Thr	GAC Asp	TTT Phe	CAA Gln	TCC Ser 755	2493
15			AGT Ser														2541
20	Thr	Pro	ACT Thr	Ser 775	Lys	Asp	Val	Leu	Ser 780	Asn	Leu	Val	Met	Ile 785	Ser	Arg	2589
25			GAA Glu 790														2637
			GAT Asp														2685
30	GAT Asp 820	Val	TGT Cys	GCT Ala	TTA Leu	AAT Asn 825	GAA Glu	AAT Asn	TAT Tyr	AAA Lys	AAC Asn 830	GTT Val	GAG Glu	CTG Leu	TTG Leu	CCA Pro 835	2733
35			AAA Lys														2781
40			AAC Asn														2829
45	ACT Thr	TCA Ser	ATT Ile 870	TCA Ser	AAA Lys	ATA Ile	ACT Thr	GTC Val 875	AAT Asn	CCA Pro	GAC Asp	TCT Ser	GAA Glu 880	GAA Glu	CTT Leu	TTC Phe	2877
			AAT Asn					Val									2925
50		Leu	GCT Ala				Thr					Glu					2973
55						Ile					Thr					GGA Gly	3021
60					Lys					Val					Asp	TTG Leu	3069
	GTT	TAT	GTT	CTT	GCA	GAG	GAG	AAC	. AAA	. AAT	' AGT	GTA	AAG	CAG	CAT	ATA	3117

	Val	Tyr	Val 950	Leu	Ala	Glu	Glu	Asn 955	Lys	Asn	Ser	Val	Lys 960	Gln	His	Ile	
5		ATG Met 965															3165
10		AAA Lys															3213
15		TTA Leu		Pro					Ser					Phe			3261
20		TCA Ser	Asn					Leu					Ile				3309
20		ATG Met					Ile					Pro					3357
25	Cys	GTT Val 1045				Asn					Asp						3405
30		AAG Lys			Ser					Ser					Ser		3453
35		GTT Val		Ser					Ser					Gln			3501
4.0		TCC Ser	Lys					Ser					Thr				3549
40		GCA Ala					Leu					Glu					3597
45		TTT Phe	Glu			Gln					Ser						3645
50		ACA Thr			Val		Glu			Met					Thr		3693
55		GAG Glu		Cys					Leu					Asn			3741
60		ATT											Gly				3789
60		AAA Lys															3837

1190 1195 1200

5	GCT TCT GGT TAT TTA ACA GAT GAA AAT GAA GTG GGG TTT AGG GGC TTT Ala Ser Gly Tyr Leu Thr Asp Glu Asn Glu Val Gly Phe Arg Gly Phe 1205 1210 1215	3885
10	TAT TCT GCT CAT GGC ACA AAA CTG AAT GTT TCT ACT GAA GCT CTG CAA Tyr Ser Ala His Gly Thr Lys Leu Asn Val Ser Thr Glu Ala Leu Gln 1220 1225 1230 1235	3933
1.5	AAA GCT GTG AAA CTG TTT AGT GAT ATT GAG AAT ATT AGT GAG GAA ACT Lys Ala Val Lys Leu Phe Ser Asp Ile Glu Asn Ile Ser Glu Glu Thr 1240 1245 1250	3981
15	TCT GCA GAG GTA CAT CCA ATA AGT TTA TCT TCA AGT AAA TGT CAT GAT Ser Ala Glu Val His Pro Ile Ser Leu Ser Ser Lys Cys His Asp 1255 1260 1265	4029
20	TCT GTT GTT TCA ATG TTT AAG ATA GAA AAT CAT AAT GAT AAA ACT GTA Ser Val Val Ser Met Phe Lys Ile Glu Asn His Asn Asp Lys Thr Val 1270 1275 1280	4077
25	AGT GAA AAA AAT AAA TGC CAA CTG ATA TTA CAA AAT AAT ATT GAA Ser Glu Lys Asn Asn Lys Cys Gln Leu Ile Leu Gln Asn Asn Ile Glu 1285 1290 1295	4125
30	ATG ACT ACT GGC ACT TTT GTT GAA GAA ATT ACT GAA AAT TAC AAG AGA Met Thr Thr Gly Thr Phe Val Glu Glu Ile Thr Glu Asn Tyr Lys Arg 1300 1305 1310 1315	4173
	AAT ACT GAA AAT GAA GAT AAC AAA TAT ACT GCT GCC AGT AGA AAT TCT Asn Thr Glu Asn Glu Asp Asn Lys Tyr Thr Ala Ala Ser Arg Asn Ser 1320 1325 1330	4221
35	CAT AAC TTA GAA TTT GAT GGC AGT GAT TCA AGT AAA AAT GAT ACT GTT His Asn Leu Glu Phe Asp Gly Ser Asp Ser Ser Lys Asn Asp Thr Val 1335 1340 1345	4269
40	TGT ATT CAT AAA GAT GAA ACG GAC TTG CTA TTT ACT GAT CAG CAC AAC Cys Ile His Lys Asp Glu Thr Asp Leu Leu Phe Thr Asp Gln His Asn 1350 1355 1360	4317
45	ATA TGT CTT AAA TTA TCT GGC CAG TTT ATG AAG GAG GGA AAC ACT CAG  Ile Cys Leu Lys Leu Ser Gly Gln Phe Met Lys Glu Gly Asn Thr Gln  1365 1370 1375	4365
50	ATT AAA GAA GAT TTG TCA GAT TTA ACT TTT TTG GAA GTT GCG AAA GCT Ile Lys Glu Asp Leu Ser Asp Leu Thr Phe Leu Glu Val Ala Lys Ala 1380 1385 1390 1395	4413
5.5	CAA GAA GCA TGT CAT GGT AAT ACT TCA AAT AAA GAA CAG TTA ACT GCT Gln Glu Ala Cys His Gly Asn Thr Ser Asn Lys Glu Gln Leu Thr Ala 1400 1405 1410	4461
55	ACT AAA ACG GAG CAA AAT ATA AAA GAT TTT GAG ACT TCT GAT ACA TTT Thr Lys Thr Glu Gln Asn Ile Lys Asp Phe Glu Thr Ser Asp Thr Phe 1415 1420 1425	4509
60	TTT CAG ACT GCA AGT GGG AAA AAT ATT AGT GTC GCC AAA GAG TCA TTT Phe Gln Thr Ala Ser Gly Lys Asn Ile Ser Val Ala Lys Glu Ser Phe 1430 1435 1440	4557

5	AAT AAA Asn Lys 1445		Phe					Pro					4605
10	TTT TCC Phe Ser 1460						Asp					Lys	4653
	GAC ATT Asp Ile		Glu			Asp					Lys		4701
15	AAA GAA Lys Glu	Ser			Thr					Val			4749
20	GGA CAA Gly Gln			Glu					Pro				4797
25	TTT CAT Phe His 1525	Thr	Gly					Ile					4845
30	GAC AAA Asp Lys 1540						Lys					Ser	4893
30	ATC ACC		His			Ala					Tyr		4941
35	GCC TGT Ala Cys	Lys			Ala					Glu			4989
40	GCC CCA Ala Pro			Met					Asn				5037
45	CTT GTT Leu Val 1605	Ser	Thr					Lys					5085
50	TTA TGT Leu Cys 1620						Thr					Phe	5133
30	AAA GTT Lys Val		Glu			Glu					Lys		5181
55	GCA ACT Ala Thr	Cys			Ser					Ile			5229
60	GCC TTA Ala Lei			Ser					Thr				5277

5	ACT TCA TTA CTT GAA GCA AAA AAA TGG CTT AGA GAA GGA ATA TTT GAT Thr Ser Leu Leu Glu Ala Lys Lys Trp Leu Arg Glu Gly Ile Phe Asp 1685 1690 1695	5325
5	GGT CAA CCA GAA AGA ATA AAT ACT GCA GAT TAT GTA GGA AAT TAT TTG Gly Gln Pro Glu Arg Ile Asn Thr Ala Asp Tyr Val Gly Asn Tyr Leu 1700 1705 1710 1715	5373
10	TAT GAA AAT AAT TCA AAC AGT ACT ATA GCT GAA AAT GAC AAA AAT CAT Tyr Glu Asn Asn Ser Asn Ser Thr Ile Ala Glu Asn Asp Lys Asn His 1720 1725 1730	5421
15	CTC TCC GAA AAA CAA GAT ACT TAT TTA AGT AAC AGT AGC ATG TCT AAC Leu Ser Glu Lys Gln Asp Thr Tyr Leu Ser Asn Ser Ser Met Ser Asn 1735 1740 1745	5469
20	AGC TAT TCC TAC CAT TCT GAT GAG GTA TAT AAT GAT TCA GGA TAT CTC Ser Tyr Ser Tyr His Ser Asp Glu Val Tyr Asn Asp Ser Gly Tyr Leu 1750 1755 1760	5517
ي د	TCA AAA AAT AAA CTT GAT TCT GGT ATT GAG CCA GTA TTG AAG AAT GTT Ser Lys Asn Lys Leu Asp Ser Gly Ile Glu Pro Val Leu Lys Asn Val 1765 1770 1775	5565
25	GAA GAT CAA AAA AAC ACT AGT TTT TCC AAA GTA ATA TCC AAT GTA AAA Glu Asp Gln Lys Asn Thr Ser Phe Ser Lys Val Ile Ser Asn Val Lys 1780 1785 1790 1795	5613
30	GAT GCA AAT GCA TAC CCA CAA ACT GTA AAT GAA GAT ATT TGC GTT GAG Asp Ala Asn Ala Tyr Pro Gln Thr Val Asn Glu Asp Ile Cys Val Glu 1800 1805 1810	5661
35	GAA CTT GTG ACT AGC TCT TCA CCC TGC AAA AAT AAA AAT GCA GCC ATT Glu Leu Val Thr Ser Ser Ser Pro Cys Lys Asn Lys Asn Ala Ala Ile 1815 1820 1825	5709
40	AAA TTG TCC ATA TCT AAT AGT AAT AAT TTT GAG GTA GGG CCA CCT GCA Lys Leu Ser Ile Ser Asn Ser Asn Asn Phe Glu Val Gly Pro Pro Ala 1830 1835 1840	5757
4 =	TTT AGG ATA GCC AGT GGT AAA ATC GTT TGT GTT TCA CAT GAA ACA ATT Phe Arg Ile Ala Ser Gly Lys Ile Val Cys Val Ser His Glu Thr Ile 1845 1850 1855	5805
45	AAA AAA GTG AAA GAC ATA TTT ACA GAC AGT TTC AGT AAA GTA ATT AAG Lys Lys Val Lys Asp Ile Phe Thr Asp Ser Phe Ser Lys Val Ile Lys 1860 1865 1870 1875	5853
50	GAA AAC AAC GAG AAT AAA TCA AAA ATT TGC CAA ACG AAA ATT ATG GCA Glu Asn Asn Glu Asn Lys Ser Lys Ile Cys Gln Thr Lys Ile Met Ala 1880 1885 1890	5901
55	GGT TGT TAC GAG GCA TTG GAT GAT TCA GAG GAT ATT CTT CAT AAC TCT Gly Cys Tyr Glu Ala Leu Asp Asp Ser Glu Asp Ile Leu His Asn Ser 1895 1900 1905	5949
60	CTA GAT AAT GAT GAA TGT AGC ACG CAT TCA CAT AAG GTT TTT GCT GAC Leu Asp Asn Asp Glu Cys Ser Thr His Ser His Lys Val Phe Ala Asp 1910 1915 1920	5997
	ATT CAG AGT GAA GAA ATT TTA CAA CAT AAC CAA AAT ATG TCT GGA TTG	6045

	Ile Gln 1925	Ser Glu		Leu Glr 1930	n His A		Asn Met S 935	Ser Gly	Leu
5	GAG AAA Glu Lys 1940	GTT TCT Val Ser	AAA ATA Lys Ile 1945	TCA CCT Ser Pro	TGT G.	AT GTT A sp Val S 1950	AGT TTG C Ser Leu C	3lu Thr	TCA 6093 Ser 955
10	GAT ATA Asp Ile	Cys Lys				eu His I			
15	GCA AAT Ala Asn						Gly Lys S		
20	GTA TCA Val Ser 1	GAT GCT Asp Ala	TCA TTA Ser Leu	CAA AAG Gln Asr 1995	n Ala A	.GA CAA ( .rg Gln \	GTG TTT T Val Phe S 2000	TCT GAA Ser Glu	ATA 6237 Ile
20	GAA GAT Glu Asp 2005	AGT ACC Ser Thr	Lys Gln	GTC TT Val Phe 2010	r TCC A e Ser L	ys Val I	TTG TTT A Leu Phe D 015	AAA AGT Lys Ser	AAC 6285 Asn
25	GAA CAT Glu His 2020	TCA GAC Ser Asp	CAG CTC Gln Leu 2025	ACA AGA	A GAA G g Glu G	SAA AAT A Slu Asn 1 2030	ACT GCT A	Ile Arg	ACT 6333 Thr 1035
30	CCA GAA Pro Glu	His Leu	ATA TCC Ile Ser 2040	CAA AA Gln Ly	s Gly P	TTT TCA T Phe Ser T 045	TAT AAT ( Tyr Asn '	GTG GTA Val Val 2050	AAT 6381 Asn
35			Ser Gly				GGA AAG Gly Lys 2		
40	Ile Leu	GAA AGT Glu Ser 2070	TCC TTA Ser Leu	CAC AA His Ly 207	s Val I	AAG GGA ( Lys Gly )	GTG TTA Val Leu 2080	GAG GAA Glu Glu	TTT 6477 Phe
10						His Tyr	TCA CCT Ser Pro 095		
45	CAA AAT Gln Asn 2100	GTA TCA Val Ser	AAA ATA Lys Ile 2105	Leu Pr	T CGT ( o Arg \	GTT GAT Val Asp 2110	AAG AGA Lys Arg	Asn Pro	GAG 6573 Glu 2115
50					u Lys T		AGT AAA Ser Lys		
55			Leu Asr				TCA GAA Ser Glu 2		
60	Ser Ile				u Ser (		CAA CAA Gln Gln 2160		
00							GAG AAC Glu Asn		

2165 2170 2175

5				Lys Asn		ATG GAA ATT Met Glu Ile	
10	Glu Thr					AAT ATA GAA Asn Ile Glu 2210	
1.5						GAA ACA GAA Glu Thr Glu 2225	
15				Glu Asp	Asp Glu	CTG ACA GAT Leu Thr Asp 2240	
20	Pro Ser	His Ala				TGT CCC GAA Cys Pro Glu	
25			Asn Ser	Arg Ile		AGA AGA GGA Arg Arg Gly	
30	lle Leu					AAC TTA TTA Asn Leu Leu 2290	Asn
						TTA AAG GCT Leu Lys Ala 2305	
35				e Lys Asp	Arg Arg	TTG TTT ATO Leu Phe Met 2320	
40	Ser Leu					CGC ACA ACT	
45			Asn Pro	Asn Phe		CCT GGT CAP	_
50					Leu Thr	TTG GAA AAA Leu Glu Lys 2370	s Ser
		Ala Val				CAA GTT TCT Gln Val Ser 2385	
55				s Leu Ile	Thr Thr	GGC AGA CCA Gly Arg Pro 2400	
60	l Phe Val					TTT CAC AGA	

5	GAA G Glu G 2420				Arg					Glu					Lys		7533
10	AAC A Asn I			Gly					Asp					Ile			7581
10	AAT G Asn G		Ile					Lys					Gln				7629
15	GTA A	Chr					Glu					Asp					7677
20	CTT C Leu C					Asp					Arg						7725
25	AGG ( Arg ( 2500				Phe					Ser					Lys		7773
2.0	TCC A			Pro					Lys					Gly			7821
30	CCC T		Ala					Gln					Gly				7869
35	CAT T	Cys					Ser					Ser					7917
40	ACT ( Thr (					Gly					Trp				_	_	7965
45	CAG 5 Gln 1 2580				Gly					Pro					Lys		8013
50	GGA Z			Glu					Leu					Gly			8061
30	CCA /		Leu					Trp					Tyr				8109
55	ATA '	Trp					Met					Pro				_	8157
60	AAT . Asn . 2					Pro					Leu						8205

5					Ile					Arg					AAG Lys 2		8253
3				Asp					Lys					Cys	GTT Val 2690		8301
10			Ile					Asn					Ser		AAT Asn		8349
15		Ser					Gln					Ile			ACA Thr		8397
20	Gly					Lys					Pro				GCT Ala	_	8445
25		Lys			Arg					Gln					CAT His		8493
				Val					Ala					Glu	GCC Ala 2770		8541
30			Leu					Ser					Arg		GCT Ala		8589
35		Tyr					Phe					Arg			CCT Pro		8637
40	Pro		Ser			Phe					Asn				GTT Val		8685
45		Ile			Arg					Gln					ACA Thr		8733
				Tyr					Glu					Lys	GAA Glu 2850		8781
50			Tyr					Gln					Ala		TTC Phe		8829
55				Glu			Glu					Asn			AAA Lys		8877
60			Pro			Ala					Gln				TTG Leu		8925
	GAT	GGT	GCA	GAG	CTT	TAT	GAA	GCA	GTG	AAG	AAT	GCA	GCA	GAC	CCA	GCT	8973

	Asp 2900	Gly	Ala	Glu		Tyr :905	Glu	Ala	Val	-	Asn 910	Ala	Ala	Asp		Ala 915	
5	TAC Tyr			Gly					Glu					Leu	AAT Asn 2930		9021
10	CAC		Gln					Lys					Ile				9069
15		Arg					Ser					Glu			TTA Leu		9117
20	Arg					Val					Ile				TCA Ser		9165
20					Ser					Ile					TCA Ser		9213
25				Leu					Lys					Tyr	CAT His 3010		9261
30			Ser					Lys					Asn		CAG Gln		9309
35		Ala					Gln					Pro			GAT Asp		9357
40	Ile					Tyr					Pro				AGC Ser		9405
40					Asp					Cys					CTA Leu		9453
45				Val					Lys					Pro	TTC Phe 3090		9501
50			Ser					Asn					Lys		TGG Trp		9549
55		Leu					Ile					Leu			GCA Ala		9597
60	Asn					Pro					Gly				TTA Leu		9645
50															CAC His		9693

CTC AGA CTG AAA CGA CGT TGT ACT ACA TCT CTG ATC AAA GAA CAG GAG

Leu Arg Leu Lys Arg Arg Cys Thr Thr Ser Leu Ile Lys Glu Glu Glu

AGT TCC CAG GCC AGT ACG GAA GAA TGT GAG AAA AAT AAG CAG GAC ACA Ser Ser Gln Ala Ser Thr Glu Glu Cys Glu Lys Asn Lys Gln Asp Thr ATT ACA ACT AAA AAA TAT ATC TAA Ile Thr Thr Lys Lys Tyr Ile (2) INFORMATION FOR SEQ ID NO:13: (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 3418 amino acids (B) TYPE: amino acid (C) STRANDEDNESS: single (D) TOPOLOGY: linear (ii) MOLECULE TYPE: protein (v) FRAGMENT TYPE: internal (xi) SEQUENCE DESCRIPTION: SEQ ID NO:13: Met Pro Ile Gly Ser Lys Glu Arg Pro Thr Phe Phe Glu Ile Phe Lys Thr Arg Cys Asn Lys Ala Asp Leu Gly Pro Ile Ser Leu Asn Trp Phe Glu Glu Leu Ser Ser Glu Ala Pro Pro Tyr Asn Ser Glu Pro Ala Glu Glu Ser Glu His Lys Asn Asn Tyr Glu Pro Asn Leu Phe Lys Thr Pro Gln Arg Lys Pro Ser Tyr Asn Gln Leu Ala Ser Thr Pro Ile Ile Phe Lys Glu Gln Gly Leu Thr Leu Pro Leu Tyr Gln Ser Pro Val Lys Glu Leu Asp Lys Phe Lys Leu Asp Leu Gly Arg Asn Val Pro Asn Ser Arg His Lys Ser Leu Arg Thr Val Lys Thr Lys Met Asp Gln Ala Asp Asp Val Ser Cys Pro Leu Leu Asn Ser Cys Leu Ser Glu Ser Pro Val Val Leu Gln Cys Thr His Val Thr Pro Gln Arg Asp Lys Ser Val Val Cys Gly Ser Leu Phe His Thr Pro Lys Phe Val Lys Gly Arg Gln Thr Pro Lys His Ile Ser Glu Ser Leu Gly Ala Glu Val Asp Pro Asp Met Ser Trp Ser Ser Ser Leu Ala Thr Pro Pro Thr Leu Ser Ser Thr Val Leu Ile Val Arg Asn Glu Glu Ala Ser Glu Thr Val Phe Pro His Asp Thr Thr Ala Asn Val Lys Ser Tyr Phe Ser Asn His Asp Glu Ser Leu Lys Lys Asn Asp Arg Phe Ile Ala Ser Val Thr Asp Ser Glu Asn Thr Asn Gln Arg Glu Ala Ala Ser His Gly Phe Gly Lys Thr Ser Gly Asn Ser Phe Lys Val Asn Ser Cys Lys Asp His Ile Gly Lys Ser Met Pro His Val Leu Glu Asp Glu Val Tyr Glu Thr Val Val Asp Thr Ser Glu

	305	Asp				310					315					320
5		Lys			325					330					335	
J		Ala		340					345					350		
		Phe	355					360					365			
10		Val 370					375					380				
	385	Glu				390					395					400
15		Gly			405					410					415	
		Ser		420					425					430		
2.0		Lys Ser	435					440					445			
20		450 Asn					455					460				
	465	Leu				470					475					480
25		Phe			485					490					495	
		Glu		500					505					510		
30	-	Lys	515					520					525			
	Val	530 Cys	Ser	Gln	Lys	Glu	535 Asp	Ser	Leu	Cys		540 Asn	Leu	Ile	Asp	Asn
	545 Gly	Ser	Trp	Pro		550 Thr	Thr	Thr	Gln		555 Ser	Val	Ala	Leu	Lys 575	560 Asn
35	Ala	Gly	Leu		565 Ser	Thr	Leu	Lys	Lys 585	570 Lys	Thr	Asn	Lys	Phe 590		Tyr
	Ala	Ile	His 595	580 Asp	Glu	Thr	Ser	Tyr 600	Lys	Gly	Lys	Lys	Ile 605		Lys	Asp
40	Gln	Lys 610	Ser	Glu	Leu	Ile	Asn 615	Cys		Ala	Gln	Phe 620	Glu	Ala	Asn	Ala
	Phe 625	Glu	Ala	Pro	Leu	Thr 630		Ala	Asn	Ala	Asp 635	Ser	Gly	Leu	Leu	His 640
45	Ser	Ser			645					650					655	
				660					665					670		Arg -
			675					680					685			Tyr
50	_	690					695	;				700				Pro
	705	,				710	1				715					720 Ala
55					725					730	)				735	Ala Asp
				740					745	5				750	•	Thr
60			755					760	)				765	5		Met
5.0		770	1				775	5				780	)			: Gly
			_	. 4	-			-	-			_				

	785					790					795					800
					805					810					815	
5				820					825					Asn 830		
			835					840					845	Ser		
10		850					855					860		Lys		
	865					870					875			Asp		880
					885					890				Ile	895	
15				900					905					His 910		
	_		915					920					925	Thr		
20		930	_				935					940		Ser		
	945	_			_	950					955			Ser		960
					965					970				Asp	975	
25				980					985					Met 990		
			995					1000	)				1005			
30		1010	)				1015	5				1020	)	His		
	1025	5				1030	)				1035	5		Tyr		104
					1045	5				1050	)			Asp	1055	5
35	_	_		1060	)				1069	5				Ala 1070	C	
			1079	5				1080	0				108			
40		109	0				109	5	Pne	Asn	ser	1100		Asn	Leu	THE
	1109		GIn	LVS	Δla			1	~ 1	<b>-</b>	<b>a</b>	m1		T	a1	~1
				_		1110	)				1115	5	Ile	Leu		112
		Gly		Gln	Phe 112!	1110 Glu 5	) Phe	Thr	Gln	Phe	111! Arg	Lys	Ile Pro	Ser	Tyr 1135	112 Ile
45	Leu	Gly Gln	Lys	Gln Ser	Phe 112! Thr	1110 Glu 5 Phe	Phe Glu	Thr Val	Gln Pro	Phe 1130 Glu	111! Arg ) Asn	Lys Gln	Ile Pro Met	Ser Thr 1150	Tyr 1135 Ile 0	112 Ile 5 Leu
45	Leu Lys	Gly Gln Thr	Lys Thr 115	Gln Ser 1140 Ser	Phe 112! Thr ) Glu	1110 Glu 5 Phe Glu	Phe Glu Cys	Thr Val Arg	Gln Pro 114! Asp	Phe 1130 Glu 5 Ala	111! Arg ) Asn Asp	Lys Gln Leu	Ile Pro Met His 116	Ser Thr 1150 Val 5	Tyr 1135 Ile O Ile	112 Ile 5 Leu Met
45 50	Leu Lys Asn	Gly Gln Thr Ala	Lys Thr 115: Pro	Gln Ser 1140 Ser 5	Phe 112! Thr ) Glu Ile	Glu Fhe Glu Glu	Phe Glu Cys Gln 117	Thr Val Arg 116 Val Val	Gln Pro 114! Asp 0 Asp	Phe 1130 Glu Ala Ser	Arg Asn Asp Ser	Lys Leu Lys 118	Ile Pro Met His 116 Gln	Ser Thr 1150 Val 5 Phe	Tyr 1139 Ile O Ile Glu	112 Ile E Leu Met
	Leu Lys Asn Thr	Gly Gln Thr Ala 117 Val	Lys Thr 115: Pro 0 Glu	Gln Ser 1140 Ser Ser Tle	Phe 112! Thr ) Glu Ile	Glu  Phe  Glu  Gly  Arg  119	Phe Glu Cys Gln 1179 Lys	Thr Val Arg 116 Val 5	Gln Pro 114! Asp O Asp	Phe 1130 Glu Ala Ser	Arg Asn Asp Ser Leu 119	Lys Gln Leu Lys 1180 Leu	Pro Met His 116 Gln O Lys	Ser Thr 1150 Val 5 Phe Asn	Tyr 1139 Ile O Ile Glu Asp	112 Ile E Leu Met Gly Cys 120
50	Leu Lys Asn Thr 118	Gly Gln Thr Ala 117 Val 5 Lys	Lys Thr 115: Pro Glu Ser	Gln Ser 1140 Ser 5 Ser Ile Ala	Phe 112! Thr Glu Ile Lys Ser 120!	Glu  Phe  Glu  Gly  Arg  Gly  Gly	Phe Glu Cys Gln 1177 Lys 0	Thr Val Arg 116 Val Fhe Leu	Gln Pro 1149 Asp O Asp Ala Thr	Phe 1130 Glu 5 Ala Ser Gly Asp 1210	Arg Asn Asp Ser Leu 119 Glu	Lys Gln Leu Lys 1180 Leu Asn	Pro Met His 116 Gln Lys Glu	Ser Thr 1150 Val 5 Phe Asn Val	Tyr 1135 Ile 0 Ile Glu Asp Gly 1215	112 Ile 5 Leu Met Gly Cys 120 Phe
	Leu Lys Asn Thr 118 Asn	Gly Gln Thr Ala 117 Val 5 Lys	Lys Thr 1155 Pro Glu Ser	Gln Ser 1140 Ser 5 Ser Ile Ala Tyr 1220	Phe 1129 Thr Glu Ile Lys Ser 1209 Ser	Glu Glu Glu Gly Arg Gly Gly Ala	Phe Glu Cys Gln 117: Lys Tyr His	Thr Val Arg 116 Val 5 Phe Leu Gly	Gln Pro 1149 Asp O Asp Ala Thr	Phe 1130 Glu 5 Ala Ser Gly Asp 1210 Lys	Asn Asp Ser Leu 119 Glu Leu	Lys Leu Lys 1180 Leu Asn	Pro Met His 116 Gln C Lys Glu Val	Ser Thr 1150 Val 5 Phe Asn Val Ser 123	Tyr 113: Ile 0 Ile Glu Asp Gly 121: Thr	112 Ile 5 Leu Met Gly Cys 120 Phe 5 Glu
50	Leu Lys Asn Thr 118 Asn Arg	Gly Gln Thr Ala 117 Val 5 Lys Gly Leu	Lys Thr 1155 Pro Glu Ser Phe Gln 123	Gln Ser 1140 Ser 5 Ser Ile Ala Tyr 1220 Lys 5	Phe 1129 Thr Glu Ile Lys Ser 1200 Ser	Glu Glu Glu Gly Arg 119 Gly Arg Arg Arg Arg Arg	Phe Glu Cys Gln 117: Lys Tyr His	Thr Val Arg 116 Val 5 Phe Leu Gly Leu 124	Gln Pro 114! Asp O Asp Ala Thr Thr 122 Phe O	Phe 1130 Glu 5 Ala Ser Gly Asp 1210 Lys 5	Asp Asp Ser Leu 119 Glu Leu Asp	Lys Lys Lys 1180 Leu Asn Asn	Pro Met His 116 Gln C Lys Glu Val Glu 124	Ser Thr 1150 Val 5 Phe Asn Val Ser 1230 Asn 5	Tyr 113: Ile 0 Ile Glu Asp Gly 121: Thr 0	112 Ile 5 Leu Met Gly Cys 120 Phe 5 Glu Ser
50	Leu Lys Asn Thr 1188 Asn Arg Ala Glu	Gly Gln Thr Ala 117 Val 5 Lys Gly Leu Glu 125	Lys Thr 1155 Pro O Glu Ser Phe Gln 123 Thr	Gln Ser 1140 Ser 5 Ser Ile Ala Tyr 1220 Lys 5 Ser	Phe 1129 Thr Glu Ile Lys Ser 1200 Ser Ala	Glu Glu Glu Gly Arg Gly Gly Arg Gly Gly Gly Glu Glu Glu	Phe Glu Cys Gln 117: Lys Tyr His Lys Val 125	Thr Val Arg 116 Val 5 Phe Leu Gly Leu 124 His 5	Gln Pro 114! Asp O Asp Ala Thr Thr 122 Phe O Pro	Phe 1130 Glu 5 Ala Ser Gly Asp 1210 Lys 5 Ser	Asn Asp Ser Leu 1199 Glu Leu Asp	Leu Lys 1180 Leu Asn Asn Ile Leu 126	Pro Met His 116 Gln C Lys Glu Val Glu 124 Ser	Ser Thr 1150 Val 5 Phe Asn Val Ser 1230 Asn	Tyr 113: Ile 0 Ile Glu Asp Gly 121: Thr 0 Ile Ser	112 Ile Leu Met Gly Cys 120 Phe 5 Glu Ser Lys

	Lys	Thr	Val	Ser	Glu 1285		Asn	Asn	Lys	Cys 1290		Leu	Ile	Leu	Gln 1295	
5	Asn	Ile	Glu	Met 1300	Thr		Gly	Thr	Phe 1305		Glu	Glu	Ile	Thr 1310		Asn
	-	Lys	1315	;				1320	)				1325	5		
	_	Asn 1330	)				1335	5				1340	)			
10	1345					1350	)				1355	,				136
		His			1365	;				1370	)				1375	j
15		Thr		1380	)				1385	5				1390	)	
		Lys	1395	5				1400	)				1405	5		
2.0		Thr	)				1415	5				1420	)			
20	142					1430	)				1435	5				144
		Ser			1445	;				1450	)				1455	5
25		His		1460	)				1465	5				1470	)	
		Lys	1475	5				1480	)				1485	5		
2.0	_	Ile 1490	)				1499	5				1500	)			
30	150			_		1510	)				1515	5				152
		Leu Ser			1525	5				1530	)				1535	5
35		Ser		1540	) _				1549	5				1550	)	
		Arg	1555	5				1560	)				156	5		
40	-	1570 Thr	0		_	_	157	5				1580	0			
10	158					1590	0				1595	5				160
	_	Asp			1605	5				1610	C				161	5
45		Phe		1620	0				162	5				163	0	
		Ser	1635	5				1640	)				164	5		
50	_	165 Asn	0			-	165	5				1660	0			
	166					167	0				167	5				168
		Phe			1685	5				169	0				169	5
55	Asn	Tyr	Leu	170 Tyr		Asn	Asn	Ser	170 Asn		Thr	Ile	Ala	171 Glu		Asp
	Lys	Asn	171! His		Ser	Glu				Thr	Tyr				Ser	Ser
60	Met	173 Ser		Ser	Tyr				Ser	Asp				Asn	Asp	
	174 Gly	5 Tyr	Leu	Ser	Lys	175 Asn		Leu	Asp	Ser	175. Gly		Glu	Pro	Val	176 Leu

		1765		1770	1775
	1	lu Asp Gln .780	1785	5	Lys Val Ile Ser 1790
5	1795		1800		Asn Glu Asp Ile 1805
	1810		1815	1820	
10	1825	1830	)	1835	Phe Glu Val Gly 184
		1845		1850	Cys Val Ser His 1855
4.5	1	860	1865	5	Ser Phe Ser Lys 1870
15	1875		1880		Cys Gln Thr Lys 1885 Glu Asp Ile Leu
	1890		1895	1900	
20	1905	1910	)	1915	192 Asn Gln Asn Met
		1925		1930	1935 Asp Val Ser Leu
25		1940	194	5	1950 Leu His Lys Ser
25	1955		1960		1965 Ala Ser Gly Lys
	1970		1975	1980	
30	1985	1990	0	1995	200
		2005		2010	Lys Val Leu Phe 2015 Glu Asn Thr Ala
2.5		2020	202	5	2030 Phe Ser Tyr Asn
35	2035		2040		2045
	2050		2055	206	
40	2065	2070	0	2075	Lys Gly Val Leu 208
		2085		2090	His Tyr Ser Pro 2095
4.5		2100	210	5	Val Asp Lys Arg 2110 The Cys See Lys
45	2115		2120		Thr Cys Ser Lys 2125 Gly Ser Ser Gly
	2130		2135	214	Gly Ser Ser Glu O Gln Phe Gln Gln
50	2145	215	0	2155	216
	_	2165		2170	Leu Val Glu Asn 2175
		2180	218	5	Asn Val Lys Met 2190 Val Lys Thr Asn
55	2195		2200		2205 Asn Tyr Phe Glu
	2210	_	2215	222	
60	2225	223	0	2235	Leu Phe Thr Cys
	III Aby bei	2245	JOI HID AIG	2250	2255

		Pro	Glu	Asn	Glu 2260		Met	Val	Leu	Ser 2265		Ser	Arg	Ile	Gly 2270		Arg
Ē		Arg	Gly	Glu 2275		Leu	Ile	Leu	Val 2280	Gly		Pro	Ser	Ile 2285	Lys	Arg	Asn
			2290	)				2295	; ;				2300	)	Lys		
		2305	5		_		2310	)				2315	5		Arg		232
1(						2325					2330	)			Pro	2335	5
				_	2340	)				2345	5				Thr 2350	)	
15		_		2355	5				2360	)				2365			
			2370	)				2375	5				2380	)	Phe		
		2385	5			-	2390	)				2395	5		Thr		240
20		_			-	2405	5				2410	)			Ser	2415	5
					2420	)	=			2425	5				Glu 2430	)	
2!	5		-	2435	5		_	_	2440	)				244			
			2450	)				2455	5				2460	)	Ser		
		246	5				2470	)	_	_		2475	5		Leu		248
31	0					2485	5				2490	)			Arg	2495	5
		_	_		2500	)				250	5				Leu 2510	)	
3	5		•	2519	5				2520	)				252			
		_	253	)				253	5				254	0	Thr		
	•	254	5				2550	C				255	5		Glu		256
4	O					2569	5				2570	)			Trp	2579	5
					2580	)				258	5				Ser 259	0	
4	5	_	_	259	5	_			260	0				260			
		_	261	o _				261	5				262	0	Asn		
5	0	262	5			_	263	0				263	5		Phe Leu		264
5	U					264	5				265	0			Ser	265	5
		_	_		266	0				266	5				267 Leu	0	
5	5	-	-	267	5			_	268	0				268			
		_	269	0	_			269	5				270	0	Ile		
6	0	270	5	_			271	0				271	5		Pro		272
J	-					272	5				273	0			Lys	273	5
		лeи	AIG	vai	пeп	пys	WOII	$\Delta T \lambda$	AT 9	шeu	TIIT	v ct T	$a \perp b$	0111	-ys		110

				2740					2745					2750		
	Leu	His	Gly 2755	Ala		Leu			Ser		Asp	Ala		Thr		Leu
5	Glu	Ala 2770		Glu	Ser	Leu	Met 2775		Lys	Ile	Ser	Ala 2780		Ser	Thr	Arg
	Pro 2785		Arg	Trp	Tyr	Thr 2790		Leu	Gly	Phe	Phe 2795		Asp	Pro	Arg	Pro 280
10	Phe	Pro	Leu	Pro	Leu 2805		Ser	Leu	Phe	Ser 2810		Gly	Gly		Val 2815	
	Cys	Val	Asp	Val 2820		Ile	Gln	Arg	Ala 2825		Pro	Ile	Gln	Trp 2830		Glu
			2835					2840	)				2845	5		
15	_	2850	)	Ala			2855	,				2860	)			
	2865	5		Lys		2870	)				2875	5				288
20		-		Tyr	2885	5				2890	)				2895	5
				Asp 2900	)				2905	5				2910	)	
			2915					2920	)				2925	5		
25		2930	)	His			2935	5				2940	)			
	2945	5		Ile		2950	)				2955	5				296
30	_			Arg	2965	3				2970	)				2975	5
				Lys 2980	)				2985	5				2990	)	
			299					3000	)				3005	5		
35	_	301	0	Ala			3019	5				3020	)			
	302	5		Ala		303	0				303	5				304
40		_		Ile	304	5				305	0				305	5
				Phe 3060	<b>O</b>				306	5				3070	)	
4.5	_		307					308	0				308	5		Ala
45		309	0	_			309	5				310	0			Ile
	310	5		Asp		311	0				311	5				312
50				Ala	312	5				313	0				313	5
				314	0				314	5				315	0	Asn
55	_		315	5				316	0				316	5		Leu
55		317	0		_		317	5				318	0			Ser
	318	5				319	0				319	5				320 Leu
60					320	5				321	0				321	
	1-1C C	SCI	DCI	322		Cys	Olu	110	322		<b></b>	201		323		

	Cys Met Ala Lys Arg Lys Ser Val Ser Thr Pro Val S 3235 3240	Ser Ala Gln Met 3245
5	Thr Ser Lys Ser Cys Lys Gly Glu Lys Glu Ile Asp A	
5	Cys Lys Lys Arg Arg Ala Leu Asp Phe Leu Ser Arg I	Leu Pro Leu Pro 328
	Pro Pro Val Ser Pro Ile Cys Thr Phe Val Ser Pro A	
10	Ala Phe Gln Pro Pro Arg Ser Cys Gly Thr Lys Tyr G	Glu Thr Pro Ile
	3300 3305 Lys Lys Lys Glu Leu Asn Ser Pro Gln Met Thr Pro F	
1.	Asn Glu Ile Ser Leu Leu Glu Ser Asn Ser Ile Ala A	3325 Asp Glu Glu Leu
15	3330 3335 3340 Ala Leu Ile Asn Thr Gln Ala Leu Leu Ser Gly Ser T	
	3345 3350 3355 Gln Phe Ile Ser Val Ser Glu Ser Thr Arg Thr Ala F	
20	3365 3370  Glu Asp Tyr Leu Arg Leu Lys Arg Arg Cys Thr Thr S	•
	3380 3385 Glu Gln Glu Ser Ser Gln Ala Ser Thr Glu Glu Cys G	
	Gln Asp Thr Ile Thr Thr Lys Lys Tyr Ile	3405
25	3410 3415	`
	(2) INFORMATION FOR SEQ ID NO:14:	
30	<ul><li>(i) SEQUENCE CHARACTERISTICS:</li><li>(A) LENGTH: 20 base pairs</li></ul>	
	(B) TYPE: nucleic acid	
	<ul><li>(C) STRANDEDNESS: single</li><li>(D) TOPOLOGY: linear</li></ul>	
35	(-)	
	(A) NAME/KEY: (B) LOCATION:	
	(D) OTHER INFORMATION: 2F primer	
40	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:	
	TGAGTTTTAC CTCAGTCACA	20
45	(2) INFORMATION FOR SEQ ID NO:16:	
	(i) SEQUENCE CHARACTERISTICS:	
	<ul><li>(A) LENGTH: 41 base pairs</li><li>(B) TYPE: nucleic acid</li></ul>	
50	(C) STRANDEDNESS: single (D) TOPOLOGY: linear	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:	
55	CAGGAAACAG CTATGACCCT GTGACGTACT GGGTTTTTAG C	41
	(2) INFORMATION FOR SEQ ID NO:17:	
<b>C</b> 0	(i) SEQUENCE CHARACTERISTICS:	
60	<ul><li>(A) LENGTH: 24 base pairs</li><li>(B) TYPE: nucleic acid</li></ul>	
	(C) STRANDEDNESS: single	

5	<ul><li>(A) NAME/KEY:</li><li>(B) LOCATION:</li><li>(D) OTHER INFORMATION: 3FII primer</li></ul>	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:	
10	GATCTTTAAC TGTTCTGGGT CACA	24
	(2) INFORMATION FOR SEQ ID NO:18:	
15	(i) SEQUENCE CHARACTERISTICS:  (A) LENGTH: 22 base pairs  (B) TYPE: nucleic acid  (C) STRANDEDNESS: single	
20	(D) TOPOLOGY: linear	
25	<ul><li>(A) NAME/KEY:</li><li>(B) LOCATION:</li><li>(D) OTHER INFORMATION: 3RII primer</li></ul>	
25	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:	
	CCCAGCATGA CACAATTAAT GA	22
30	(2) INFORMATION FOR SEQ ID NO:19:	
35	<ul><li>(i) SEQUENCE CHARACTERISTICS:</li><li>(A) LENGTH: 44 base pairs</li><li>(B) TYPE: nucleic acid</li><li>(C) STRANDEDNESS: single</li><li>(D) TOPOLOGY: linear</li></ul>	
40	<ul><li>(A) NAME/KEY:</li><li>(B) LOCATION:</li><li>(D) OTHER INFORMATION: 4F/M 13F primer</li></ul>	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:	
45	TGTAAAACGA CGGCCAGTAG AATGCAAATT TATAATCCAG AGTA	44
	(2) INFORMATION FOR SEQ ID NO:20:	
50	<ul><li>(i) SEQUENCE CHARACTERISTICS:</li><li>(A) LENGTH: 22 base pairs</li><li>(B) TYPE: nucleic acid</li><li>(C) STRANDEDNESS: single</li><li>(D) TOPOLOGY: linear</li></ul>	
55	/A NAME /KEY	
	(A) NAME/KEY: (B) LOCATION: (D) OTHER INFORMATION: 4R-1A primer	
60	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:	
	ATCAGATTCA TCTTTATAGA AC	22

(D) TOPOLOGY: linear

	(2) INFORMATION FOR SEQ ID NO:21:	
5	(i) SEQUENCE CHARACTERISTICS:  (A) LENGTH: 40 base pairs  (B) TYPE: nucleic acid	
1.0	<pre>(C) STRANDEDNESS: single (D) TOPOLOGY: linear</pre>	
10		
15	(A) NAME/KEY: (B) LOCATION: (D) OTHER INFORMATION: 5+6F/M13F primer	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:	4.0
	TGTAAAACGA CGGCCAGTTG TGTTGGCATT TTAAACATCA	40
20	(2) INFORMATION FOR SEQ ID NO:22:	
25	<ul><li>(i) SEQUENCE CHARACTERISTICS:</li><li>(A) LENGTH: 38 base pairs</li><li>(B) TYPE: nucleic acid</li><li>(C) STRANDEDNESS: single</li><li>(D) TOPOLOGY: linear</li></ul>	
30	<ul><li>(A) NAME/KEY:</li><li>(B) LOCATION:</li><li>(D) OTHER INFORMATION: 5+6R/M13R primer</li></ul>	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:	
35	CAGGAAACAG CTATGACCCA GGGCAAAGGT ATAACGCT	38
	(2) INFORMATION FOR SEQ ID NO:23:	
40	<ul><li>(i) SEQUENCE CHARACTERISTICS:</li><li>(A) LENGTH: 38 base pairs</li><li>(B) TYPE: nucleic acid</li><li>(C) STRANDEDNESS: single</li><li>(D) TOPOLOGY: linear</li></ul>	
45		
	<ul><li>(A) NAME/KEY:</li><li>(B) LOCATION:</li><li>(D) OTHER INFORMATION: 7F/M13F primer</li></ul>	
50	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:	
	TGTAAAACGA CGGCCAGTTA AGTGAAATAA AGAGTGAA	38
	(2) INFORMATION FOR SEQ ID NO:24:	
55	<ul><li>(i) SEQUENCE CHARACTERISTICS:</li><li>(A) LENGTH: 36 base pairs</li><li>(B) TYPE: nucleic acid</li></ul>	
60	(C) STRANDEDNESS: single (D) TOPOLOGY: linear	

5	<ul><li>(A) NAME/KEY:</li><li>(B) LOCATION:</li><li>(D) OTHER INFORMATION: 7R/M13R primer</li></ul>	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:	
	CAGGAAACAG CTATGACCAG AAGTATTAGA GATGAC	36
10	(2) INFORMATION FOR SEQ ID NO:25:	
15	<ul><li>(i) SEQUENCE CHARACTERISTICS:</li><li>(A) LENGTH: 40 base pairs</li><li>(B) TYPE: nucleic acid</li><li>(C) STRANDEDNESS: single</li><li>(D) TOPOLOGY: linear</li></ul>	
20	<ul><li>(A) NAME/KEY:</li><li>(B) LOCATION:</li><li>(D) OTHER INFORMATION: 8F/M13F primer</li></ul>	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:	
25	TGTAAAACGA CGGCCAGTGC CATATCTTAC CACCTTGTGA	40
	(2) INFORMATION FOR SEQ ID NO:26:	
30	<ul><li>(i) SEQUENCE CHARACTERISTICS:</li><li>(A) LENGTH: 22 base pairs</li><li>(B) TYPE: nucleic acid</li><li>(C) STRANDEDNESS: single</li><li>(D) TOPOLOGY: linear</li></ul>	
35	(ix) FEATURE:	
	<ul><li>(A) NAME/KEY:</li><li>(B) LOCATION:</li><li>(D) OTHER INFORMATION: 8FIA primer</li></ul>	
40	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:	
	TTGCATTCTA GTGATAATAT AC	22
45	(2) INFORMATION FOR SEQ ID NO:27:	
50	<ul><li>(i) SEQUENCE CHARACTERISTICS:</li><li>(A) LENGTH: 19 base pairs</li><li>(B) TYPE: nucleic acid</li><li>(C) STRANDEDNESS: single</li><li>(D) TOPOLOGY: linear</li></ul>	
55	<ul><li>(A) NAME/KEY:</li><li>(B) LOCATION:</li><li>(D) OTHER INFORMATION: 8RIA primer</li></ul>	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:	
60	AATTGTTAGC AATTTCAAC	19
	(2) INFORMATION FOR SEQ ID NO:28:	

5	<ul><li>(i) SEQUENCE CHARACTERISTICS:</li><li>(A) LENGTH: 40 base pairs</li><li>(B) TYPE: nucleic acid</li><li>(C) STRANDEDNESS: single</li><li>(D) TOPOLOGY: linear</li></ul>	
10	<ul><li>(A) NAME/KEY:</li><li>(B) LOCATION:</li><li>(D) OTHER INFORMATION: 9F/M13F primer</li></ul>	
15	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:28: TGTAAAACGA CGGCCAGTTG GACCTAGGTT GATTGCAGAT	40
20	<ul> <li>(2) INFORMATION FOR SEQ ID NO:29:</li> <li>(i) SEQUENCE CHARACTERISTICS: <ul> <li>(A) LENGTH: 40 base pairs</li> <li>(B) TYPE: nucleic acid</li> <li>(C) STRANDEDNESS: single</li> <li>(D) TOPOLOGY: linear</li> </ul> </li> </ul>	
23	(A) NAME/KEY:	
30	<ul><li>(B) LOCATION:</li><li>(D) OTHER INFORMATION: 9R/M13R primer</li><li>(xi) SEQUENCE DESCRIPTION: SEQ ID NO:29:</li></ul>	
	CAGGAAACAG CTATGACCTA AACTGAGATC ACGGGTGACA	40
35	(2) INFORMATION FOR SEQ ID NO:30:	
40	<ul><li>(i) SEQUENCE CHARACTERISTICS:</li><li>(A) LENGTH: 24 base pairs</li><li>(B) TYPE: nucleic acid</li><li>(C) STRANDEDNESS: single</li><li>(D) TOPOLOGY: linear</li></ul>	
45	<ul><li>(A) NAME/KEY:</li><li>(B) LOCATION:</li><li>(D) OTHER INFORMATION: 10AF primer</li></ul>	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:30:	
50	GAATAATATA AATTATATGG CTTA	24
	(2) INFORMATION FOR SEQ ID NO:31:	
55	<ul><li>(i) SEQUENCE CHARACTERISTICS:</li><li>(A) LENGTH: 37 base pairs</li><li>(B) TYPE: nucleic acid</li><li>(C) STRANDEDNESS: single</li><li>(D) TOPOLOGY: linear</li></ul>	
60	(A) NAME/KEY:	
	(B) LOCATION:	

	(D) OTHER INFORMATION: 10AR/M13R primer	
5	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:31:	
Э	CAGGAAACAG CTATGACCCC TAGTCTTGCT AGTTCTT	37
	(2) INFORMATION FOR SEQ ID NO:32:	
10	<ul><li>(i) SEQUENCE CHARACTERISTICS:</li><li>(A) LENGTH: 42 base pairs</li><li>(B) TYPE: nucleic acid</li><li>(C) STRANDEDNESS: single</li><li>(D) TOPOLOGY: linear</li></ul>	
15		
20	<ul><li>(A) NAME/KEY:</li><li>(B) LOCATION:</li><li>(D) OTHER INFORMATION: 10BF/M13F primer</li><li>(xi) SEQUENCE DESCRIPTION: SEQ ID NO:32:</li></ul>	
	TGTAAAACGA CGGCCAGTAR CTGAAGTGGA ACCAAATGAT AC	42
		42
25	(2) INFORMATION FOR SEQ ID NO:33:	
30	<ul> <li>(i) SEQUENCE CHARACTERISTICS:</li> <li>(A) LENGTH: 44 base pairs</li> <li>(B) TYPE: nucleic acid</li> <li>(C) STRANDEDNESS: single</li> <li>(D) TOPOLOGY: linear</li> </ul>	
35	<ul><li>(A) NAME/KEY:</li><li>(B) LOCATION:</li><li>(D) OTHER INFORMATION: 10BR/M13R primer</li></ul>	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:33:	
40	CAGGAAACAG CTATGACCAC GTGGCAAAGA ATTCTCTGAA GTAA	44
	(2) INFORMATION FOR SEQ ID NO:34:	
45	<ul><li>(i) SEQUENCE CHARACTERISTICS:</li><li>(A) LENGTH: 40 base pairs</li><li>(B) TYPE: nucleic acid</li><li>(C) STRANDEDNESS: single</li><li>(D) TOPOLOGY: linear</li></ul>	
50	(ix) FEATURE:	
	<ul><li>(ix) FEATURE:</li><li>(A) NAME/KEY:</li><li>(B) LOCATION:</li><li>(D) OTHER INFORMATION: 10CF/M13F primer</li></ul>	
50	(A) NAME/KEY: (B) LOCATION:	
	<ul><li>(A) NAME/KEY:</li><li>(B) LOCATION:</li><li>(D) OTHER INFORMATION: 10CF/M13F primer</li></ul>	40
	<pre>(A) NAME/KEY: (B) LOCATION: (D) OTHER INFORMATION: 10CF/M13F primer (xi) SEQUENCE DESCRIPTION: SEQ ID NO:34:</pre>	40

(A) LENGTH: 19 base pairs

	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:38:	
5	AAGAAGCAAA ATGTAATAAG GA	22
5	(2) INFORMATION FOR SEQ ID NO:39:	
10	<ul><li>(i) SEQUENCE CHARACTERISTICS:</li><li>(A) LENGTH: 22 base pairs</li><li>(B) TYPE: nucleic acid</li><li>(C) STRANDEDNESS: single</li><li>(D) TOPOLOGY: linear</li></ul>	
15	<ul><li>(A) NAME/KEY:</li><li>(B) LOCATION:</li><li>(D) OTHER INFORMATION: 11BR primer</li></ul>	
20	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:39:	
20	CATTTAAAGC ACATACATCT TG	22
	(2) INFORMATION FOR SEQ ID NO:40:	
25	<ul><li>(i) SEQUENCE CHARACTERISTICS:</li><li>(A) LENGTH: 21 base pairs</li><li>(B) TYPE: nucleic acid</li><li>(C) STRANDEDNESS: single</li></ul>	
30	(D) TOPOLOGY: linear	
35	<ul><li>(A) NAME/KEY:</li><li>(B) LOCATION:</li><li>(D) OTHER INFORMATION: 11CF primer</li><li>(xi) SEQUENCE DESCRIPTION: SEQ ID NO:40:</li></ul>	
	TCTAGAGGCA AAGAATCATA C	21
40	(2) INFORMATION FOR SEQ ID NO:41:	
45	<ul><li>(i) SEQUENCE CHARACTERISTICS:</li><li>(A) LENGTH: 22 base pairs</li><li>(B) TYPE: nucleic acid</li><li>(C) STRANDEDNESS: single</li><li>(D) TOPOLOGY: linear</li></ul>	
50	<ul><li>(A) NAME/KEY:</li><li>(B) LOCATION:</li><li>(D) OTHER INFORMATION: 11CR primer</li></ul>	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:41:	
55	CAAGATTATT CCTTTCATTA GC	22
60	<ul> <li>(2) INFORMATION FOR SEQ ID NO:42:</li> <li>(i) SEQUENCE CHARACTERISTICS:</li> <li>(A) LENGTH: 22 base pairs</li> <li>(B) TYPE: nucleic acid</li> <li>(C) STRANDEDNESS: single</li> </ul>	

5	<ul><li>(A) NAME/KEY:</li><li>(B) LOCATION:</li><li>(D) OTHER INFORMATION: 11DF primer</li></ul>	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:42:	
10	AACCAAAACA CAAATCTAAG AG	22
	(2) INFORMATION FOR SEQ ID NO:43:	
15	<ul><li>(i) SEQUENCE CHARACTERISTICS:</li><li>(A) LENGTH: 23 base pairs</li><li>(B) TYPE: nucleic acid</li><li>(C) STRANDEDNESS: single</li></ul>	
20	(D) TOPOLOGY: linear	
	(A) NAME/KEY: (B) LOCATION:	
25	(D) OTHER INFORMATION: 11DR primer	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:43:	
	GTCATTTTTA TATGCTGCTT TAC	23
30	(2) INFORMATION FOR SEQ ID NO:44:	
35	<ul><li>(i) SEQUENCE CHARACTERISTICS:</li><li>(A) LENGTH: 21 base pairs</li><li>(B) TYPE: nucleic acid</li><li>(C) STRANDEDNESS: single</li><li>(D) TOPOLOGY: linear</li></ul>	
40	<ul><li>(A) NAME/KEY:</li><li>(B) LOCATION:</li><li>(D) OTHER INFORMATION: 11EF primer</li></ul>	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:44:	
45	GGTTTTATAT GGAGACACAG G	21
	(2) INFORMATION FOR SEQ ID NO:45:	
50	<ul><li>(i) SEQUENCE CHARACTERISTICS:</li><li>(A) LENGTH: 23 base pairs</li><li>(B) TYPE: nucleic acid</li><li>(C) STRANDEDNESS: single</li><li>(D) TOPOLOGY: linear</li></ul>	
55		
	<ul><li>(A) NAME/KEY:</li><li>(B) LOCATION:</li><li>(D) OTHER INFORMATION: 11ER primer</li></ul>	
60	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:45:	
	GTATTTACAA TTTCAACACA AGC	23

(D) TOPOLOGY: linear

	(2) INFORMATION FOR SEQ ID NO:46:	
5	<ul><li>(i) SEQUENCE CHARACTERISTICS:</li><li>(A) LENGTH: 20 base pairs</li><li>(B) TYPE: nucleic acid</li><li>(C) STRANDEDNESS: single</li><li>(D) TOPOLOGY: linear</li></ul>	
10		
15	<ul><li>(A) NAME/KEY:</li><li>(B) LOCATION:</li><li>(D) OTHER INFORMATION: 11FF primer</li><li>(xi) SEQUENCE DESCRIPTION: SEQ ID NO:46:</li></ul>	
	ATCACAGTTT TGGAGGTAGC	20
20	(2) INFORMATION FOR SEQ ID NO:47:	
25	<ul><li>(i) SEQUENCE CHARACTERISTICS:</li><li>(A) LENGTH: 21 base pairs</li><li>(B) TYPE: nucleic acid</li><li>(C) STRANDEDNESS: single</li><li>(D) TOPOLOGY: linear</li></ul>	
30	<ul><li>(A) NAME/KEY:</li><li>(B) LOCATION:</li><li>(D) OTHER INFORMATION: 11FR primer</li></ul>	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:47:	
35	CTGACTTCCT GATTCTTCTA A	21
	(2) INFORMATION FOR SEQ ID NO:48:	
40	<ul><li>(i) SEQUENCE CHARACTERISTICS:</li><li>(A) LENGTH: 21 base pairs</li><li>(B) TYPE: nucleic acid</li><li>(C) STRANDEDNESS: single</li><li>(D) TOPOLOGY: linear</li></ul>	
45		
	<ul><li>(A) NAME/KEY:</li><li>(B) LOCATION:</li><li>(D) OTHER INFORMATION: 11GF primer</li></ul>	
50	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:48:	
	CTCAGATGTT ATTTTCCAAG C	21
	(2) INFORMATION FOR SEQ ID NO:49:	
55	<ul><li>(i) SEQUENCE CHARACTERISTICS:</li><li>(A) LENGTH: 21 base pairs</li><li>(B) TYPE: nucleic acid</li></ul>	
60	<pre>(C) STRANDEDNESS: single (D) TOPOLOGY: linear</pre>	
~ ~	(2) 10102011 111001	

5	<ul><li>(A) NAME/KEY:</li><li>(B) LOCATION:</li><li>(D) OTHER INFORMATION: 11GR primer</li></ul>	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:49:	
	CTGTTAAATA ACCAGAAGCA C	21
10	(2) INFORMATION FOR SEQ ID NO:50:	
15	<ul><li>(i) SEQUENCE CHARACTERISTICS:</li><li>(A) LENGTH: 18 base pairs</li><li>(B) TYPE: nucleic acid</li><li>(C) STRANDEDNESS: single</li><li>(D) TOPOLOGY: linear</li></ul>	
20	<ul><li>(A) NAME/KEY:</li><li>(B) LOCATION:</li><li>(D) OTHER INFORMATION: 11HF primer</li></ul>	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:50:	
25	AGGTAGACAG CAGCAAGC	18
	(2) INFORMATION FOR SEQ ID NO:51:	
30	<ul><li>(i) SEQUENCE CHARACTERISTICS:</li><li>(A) LENGTH: 22 base pairs</li><li>(B) TYPE: nucleic acid</li><li>(C) STRANDEDNESS: single</li><li>(D) TOPOLOGY: linear</li></ul>	
35	(ix) FEATURE:	
40	<pre>(A) NAME/KEY: None (B) LOCATION: (D) OTHER INFORMATION: 11HR primer (xi) SEQUENCE DESCRIPTION: SEQ ID NO:51:</pre>	
	GTAATATCAG TTGGCATTTA TT	22
45	(2) INFORMATION FOR SEQ ID NO:52:	
50	<ul><li>(i) SEQUENCE CHARACTERISTICS:</li><li>(A) LENGTH: 21 base pairs</li><li>(B) TYPE: nucleic acid</li><li>(C) STRANDEDNESS: single</li></ul>	
30	(D) TOPOLOGY: linear	
55	<ul><li>(A) NAME/KEY:</li><li>(B) LOCATION:</li><li>(D) OTHER INFORMATION: 111F primer</li><li>(xi) SEQUENCE DESCRIPTION: SEQ ID NO:52:</li></ul>	
60	TGCAGAGGTA CATCCAATAA G	21
	(2) INFORMATION FOR SEQ ID NO:53:	

5	<ul><li>(i) SEQUENCE CHARACTERISTICS:</li><li>(A) LENGTH: 21 base pairs</li><li>(B) TYPE: nucleic acid</li><li>(C) STRANDEDNESS: single</li><li>(D) TOPOLOGY: linear</li></ul>	
10	<ul><li>(A) NAME/KEY:</li><li>(B) LOCATION:</li><li>(D) OTHER INFORMATION: 11IR primer</li></ul>	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:53:	
15	GATCAGTAAA TAGCAAGTCC G	21
	(2) INFORMATION FOR SEQ ID NO:54:	
20	<ul><li>(i) SEQUENCE CHARACTERISTICS:</li><li>(A) LENGTH: 23 base pairs</li><li>(B) TYPE: nucleic acid</li><li>(C) STRANDEDNESS: single</li></ul>	
25	(D) TOPOLOGY: linear	
	(A) NAME/KEY:	
	(B) LOCATION: (D) OTHER INFORMATION: 11JF primer	
30	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:54:	
	TACTGAAAAT GAAGATAACA AAT	23
35	(2) INFORMATION FOR SEQ ID NO:55:	
40	<ul> <li>(i) SEQUENCE CHARACTERISTICS:</li> <li>(A) LENGTH: 22 base pairs</li> <li>(B) TYPE: nucleic acid</li> <li>(C) STRANDEDNESS: single</li> <li>(D) TOPOLOGY: linear</li> </ul>	
45	<ul><li>(A) NAME/KEY:</li><li>(B) LOCATION:</li><li>(D) OTHER INFORMATION: !!JR primer</li></ul>	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:55:	
50	ATTTTGTTCT TTCTTATGTC AG	22
	(2) INFORMATION FOR SEQ ID NO:56:	
55	<ul><li>(i) SEQUENCE CHARACTERISTICS:</li><li>(A) LENGTH: 35 base pairs</li><li>(B) TYPE: nucleic acid</li><li>(C) STRANDEDNESS: single</li><li>(D) TOPOLOGY: linear</li></ul>	
60	(A) NAME/KEY:	
	(B) LOCATION:	

	(D) OTHER INFORMATION: 11KF-M13 primer	
5	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:56:	
5	TGTAAAACGA CGGCCAGTCT ACTAAAACGG AGCAA	35
	(2) INFORMATION FOR SEQ ID NO:57:	
10	<ul><li>(i) SEQUENCE CHARACTERISTICS:</li><li>(A) LENGTH: 35 base pairs</li><li>(B) TYPE: nucleic acid</li><li>(C) STRANDEDNESS: single</li><li>(D) TOPOLOGY: linear</li></ul>	
15	,	
20	<ul><li>(A) NAME/KEY:</li><li>(B) LOCATION:</li><li>(D) OTHER INFORMATION: 11KR-M13 primer</li></ul>	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:57:	
	CAGGAAACAG CTATGACCGT ATGAAAACCC AACAG	35
25	(2) INFORMATION FOR SEQ ID NO:58:	
30	<ul><li>(i) SEQUENCE CHARACTERISTICS:</li><li>(A) LENGTH: 22 base pairs</li><li>(B) TYPE: nucleic acid</li><li>(C) STRANDEDNESS: single</li><li>(D) TOPOLOGY: linear</li></ul>	
35	<ul><li>(A) NAME/KEY:</li><li>(B) LOCATION:</li><li>(D) OTHER INFORMATION: 11LF primer</li></ul>	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:58:	
40	CACAAAATAC TGAAAGAAAG TG	22
	(2) INFORMATION FOR SEQ ID NO:59:	
45	<ul><li>(i) SEQUENCE CHARACTERISTICS:</li><li>(A) LENGTH: 19 base pairs</li><li>(B) TYPE: nucleic acid</li><li>(C) STRANDEDNESS: single</li><li>(D) TOPOLOGY: linear</li></ul>	
50	(A) NAME/KEY:	
	(B) LOCATION: (D) OTHER INFORMATION: 11LR primer	
55	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:59:	
	GGCACCACAG TCTCAATAG	19
60	(2) INFORMATION FOR SEQ ID NO:60:	
	(i) SEQUENCE CHARACTERISTICS:  (A) LENGTH: 20 base pairs	

5	<ul><li>(B) TYPE: nucleic acid</li><li>(C) STRANDEDNESS: single</li><li>(D) TOPOLOGY: linear</li></ul>	
5		
10	<ul><li>(A) NAME/KEY:</li><li>(B) LOCATION:</li><li>(D) OTHER INFORMATION: 11MF primer</li></ul>	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:60:	
	GCAAAGACCC TAAAGTACAG	20
15	(2) INFORMATION FOR SEQ ID NO:61:	
20	<ul><li>(i) SEQUENCE CHARACTERISTICS:</li><li>(A) LENGTH: 22 base pairs</li><li>(B) TYPE: nucleic acid</li><li>(C) STRANDEDNESS: single</li><li>(D) TOPOLOGY: linear</li></ul>	
25	<ul><li>(A) NAME/KEY:</li><li>(B) LOCATION:</li><li>(D) OTHER INFORMATION: 11MR primer</li></ul>	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:61:	
2.0	GLEGALLER MOGRECOTOR AG	
30	CATCAAATAT TCCTTCTCTA AG	22
	(2) INFORMATION FOR SEQ ID NO:62:	
35	<ul><li>(i) SEQUENCE CHARACTERISTICS:</li><li>(A) LENGTH: 35 base pairs</li><li>(B) TYPE: nucleic acid</li><li>(C) STRANDEDNESS: single</li><li>(D) TOPOLOGY: linear</li></ul>	
40		
10	<ul><li>(A) NAME/KEY:</li><li>(B) LOCATION:</li><li>(D) OTHER INFORMATION: 11NF-M13 primer</li></ul>	
45	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:62:	
	TGTAAAACGA CGGCCAGTGA AAATTCAGCC TTAGC	35
	(2) INFORMATION FOR SEQ ID NO:63:	
50	(i) SEQUENCE CHARACTERISTICS:	
	(A) LENGTH: 35 base pairs	
	(B) TYPE: nucleic acid (C) STRANDEDNESS: single	
55	(D) TOPOLOGY: linear	
	(A) NAME/KEY:	
60	(B) LOCATION:	
00	(D) OTHER INFORMATION: 11NR-M13 primer	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:63:	

CAGGAAACAG CTATGACCAT CAGAATGGTA GGAAT

5	(2) INFORMATION FOR SEQ ID NO:64:	
10	<ul><li>(i) SEQUENCE CHARACTERISTICS:</li><li>(A) LENGTH: 22 base pairs</li><li>(B) TYPE: nucleic acid</li><li>(C) STRANDEDNESS: single</li><li>(D) TOPOLOGY: linear</li></ul>	
15	<ul><li>(A) NAME/KEY:</li><li>(B) LOCATION:</li><li>(D) OTHER INFORMATION: 110F primer</li></ul>	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:64:	
20	GTACTATAGC TGAAAATGAC AA	22
	(2) INFORMATION FOR SEQ ID NO:65:	
25	<ul><li>(i) SEQUENCE CHARACTERISTICS:</li><li>(A) LENGTH: 20 base pairs</li><li>(B) TYPE: nucleic acid</li><li>(C) STRANDEDNESS: single</li><li>(D) TOPOLOGY: linear</li></ul>	
30	<ul><li>(A) NAME/KEY:</li><li>(B) LOCATION:</li><li>(D) OTHER INFORMATION: 110R primer</li></ul>	
35	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:65:	
	ACCACTGGCT ATCCTAAATG	20
40	<ul> <li>(2) INFORMATION FOR SEQ ID NO:66:</li> <li>(i) SEQUENCE CHARACTERISTICS: <ul> <li>(A) LENGTH: 20 base pairs</li> <li>(B) TYPE: nucleic acid</li> <li>(C) STRANDEDNESS: single</li> </ul> </li> </ul>	
45	(D) TOPOLOGY: linear	
50	<ul><li>(A) NAME/KEY:</li><li>(B) LOCATION:</li><li>(D) OTHER INFORMATION: 11PF primer</li><li>(xi) SEQUENCE DESCRIPTION: SEQ ID NO:66:</li></ul>	
55	TGAAGATATT TGCGTTGAGG  (2) INFORMATION FOR SEQ ID NO:67:	20
60	<ul><li>(i) SEQUENCE CHARACTERISTICS:</li><li>(A) LENGTH: 20 base pairs</li><li>(B) TYPE: nucleic acid</li><li>(C) STRANDEDNESS: single</li><li>(D) TOPOLOGY: linear</li></ul>	

5	<ul><li>(A) NAME/KEY:</li><li>(B) LOCATION:</li><li>(D) OTHER INFORMATION: 11PR primer</li></ul>	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:67:	
10	GTCAGCAAAA ACCTTATGTG	20
	(2) INFORMATION FOR SEQ ID NO:68:	
15	<ul><li>(i) SEQUENCE CHARACTERISTICS:</li><li>(A) LENGTH: 21 base pairs</li><li>(B) TYPE: nucleic acid</li><li>(C) STRANDEDNESS: single</li><li>(D) TOPOLOGY: linear</li></ul>	
20		
	<ul><li>(A) NAME/KEY:</li><li>(B) LOCATION:</li><li>(D) OTHER INFORMATION: 11QF primer</li></ul>	
25	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:68:	
	ACGAAAATTA TGGCAGGTTG T	21
	(2) INFORMATION FOR SEQ ID NO:69:	
30	(i) SEQUENCE CHARACTERISTICS:  (A) LENGTH: 21 base pairs	
35	<ul><li>(B) TYPE: nucleic acid</li><li>(C) STRANDEDNESS: single</li><li>(D) TOPOLOGY: linear</li></ul>	
40	<ul><li>(A) NAME/KEY:</li><li>(B) LOCATION:</li><li>(D) OTHER INFORMATION: 11QR primer</li></ul>	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:69:	
	CTTGTCTTGC GTTTTGTAAT G	21
45	(2) INFORMATION FOR SEQ ID NO:70:	
50	<ul><li>(i) SEQUENCE CHARACTERISTICS:</li><li>(A) LENGTH: 20 base pairs</li><li>(B) TYPE: nucleic acid</li><li>(C) STRANDEDNESS: single</li><li>(D) TOPOLOGY: linear</li></ul>	
55	<ul><li>(A) NAME/KEY:</li><li>(B) LOCATION:</li><li>(D) OTHER INFORMATION: 11RF primer</li></ul>	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:70:	
60	GCTTCATAAG TCAGTCTCAT	2.0

	(2) INFORMATION FOR SEQ ID NO:71:	
5	<ul><li>(i) SEQUENCE CHARACTERISTICS:</li><li>(A) LENGTH: 20 base pairs</li><li>(B) TYPE: nucleic acid</li><li>(C) STRANDEDNESS: single</li><li>(D) TOPOLOGY: linear</li></ul>	
10	<ul><li>(A) NAME/KEY:</li><li>(B) LOCATION:</li><li>(D) OTHER INFORMATION: 11RR primer</li></ul>	
15	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:71:	
	TCAAATTCCT CTAACACTCC	20
20	<ul><li>(2) INFORMATION FOR SEQ ID NO:72:</li><li>(i) SEQUENCE CHARACTERISTICS:</li><li>(A) LENGTH: 35 base pairs</li><li>(B) TYPE: nucleic acid</li></ul>	
25	(C) STRANDEDNESS: single (D) TOPOLOGY: linear	
30	<pre>(A) NAME/KEY: (B) LOCATION: (D) OTHER INFORMATION: 11SF-M13 primer (xi) SEQUENCE DESCRIPTION: SEQ ID NO:72:</pre>	
35	TGTAAAACGA CGGCCAGTTA CAGCAAGTGG AAAGC  (2) INFORMATION FOR SEQ ID NO:73:	35
40	<ul><li>(i) SEQUENCE CHARACTERISTICS:</li><li>(A) LENGTH: 37 base pairs</li><li>(B) TYPE: nucleic acid</li><li>(C) STRANDEDNESS: single</li><li>(D) TOPOLOGY: linear</li></ul>	
45	<ul><li>(A) NAME/KEY:</li><li>(B) LOCATION:</li><li>(D) OTHER INFORMATION: 11SR-M13 primer</li></ul>	
50	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:73:  CAGGAAACAG CTATGACCAA GTTTCAGTTT TACCAAT  (2) INFORMATION FOR SEQ ID NO:74:	37
55	<ul><li>(i) SEQUENCE CHARACTERISTICS:</li><li>(A) LENGTH: 22 base pairs</li><li>(B) TYPE: nucleic acid</li><li>(C) STRANDEDNESS: single</li><li>(D) TOPOLOGY: linear</li></ul>	
60	(-,	
	(A) NAME/KEY:	

	(B) LOCATION: (D) OTHER INFORMATION: 11TF primer	
5	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:74:	
	GTTCTTCAGA AAATAATCAC TC	22
1.0	(2) INFORMATION FOR SEQ ID NO:75:	
10	(i) SEQUENCE CHARACTERISTICS:  (A) LENGTH: 21 base pairs	
15	<ul><li>(B) TYPE: nucleic acid</li><li>(C) STRANDEDNESS: single</li><li>(D) TOPOLOGY: linear</li></ul>	
20	<ul><li>(A) NAME/KEY:</li><li>(B) LOCATION:</li><li>(D) OTHER INFORMATION: 11TR primer</li></ul>	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:75:	
25	TGTAAAAAGA GAATGTGTGG C	21
	(2) INFORMATION FOR SEQ ID NO:76:	
	<ul><li>(i) SEQUENCE CHARACTERISTICS:</li><li>(A) LENGTH: 39 base pairs</li></ul>	
30	<ul><li>(B) TYPE: nucleic acid</li><li>(C) STRANDEDNESS: single</li><li>(D) TOPOLOGY: linear</li></ul>	
35	<ul><li>(A) NAME/KEY:</li><li>(B) LOCATION:</li><li>(D) OTHER INFORMATION: 11UF-M13 primer</li></ul>	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:76:	
40	TGTAAAACGA CGGCCAGTAC TTTTTCTGAT GTTCCTGTG	39
	(2) INFORMATION FOR SEQ ID NO:77:	
45	<ul><li>(i) SEQUENCE CHARACTERISTICS:</li><li>(A) LENGTH: 39 base pairs</li><li>(B) TYPE: nucleic acid</li><li>(C) STRANDEDNESS: single</li></ul>	
50	(D) TOPOLOGY: linear	
	<ul><li>(A) NAME/KEY:</li><li>(B) LOCATION:</li><li>(D) OTHER INFORMATION: 11UR-M13 primer</li></ul>	
55	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:77:	
	CAGGAAACAG CTATGACCTA AAAATAGTGA TTGGCAACA	39
60	(2) INFORMATION FOR SEQ ID NO:78:	
	(i) SEQUENCE CHARACTERISTICS:	

	5	<ul><li>(A) LENGTH: 42 base pairs</li><li>(B) TYPE: nucleic acid</li><li>(C) STRANDEDNESS: single</li><li>(D) TOPOLOGY: linear</li></ul>	
	10	<ul><li>(A) NAME/KEY:</li><li>(B) LOCATION:</li><li>(D) OTHER INFORMATION: 12F/M13F primer</li><li>(xi) SEQUENCE DESCRIPTION: SEQ ID NO:78:</li></ul>	
	15	TGTAAAACGA CGGCCAGTAG TGGTGTTTTA AAGTGGTCAA AA	42
	10	(2) INFORMATION FOR SEQ ID NO:79:	
	20	<ul><li>(i) SEQUENCE CHARACTERISTICS:</li><li>(A) LENGTH: 40 base pairs</li><li>(B) TYPE: nucleic acid</li><li>(C) STRANDEDNESS: single</li><li>(D) TOPOLOGY: linear</li></ul>	
since had a since down south that	25	<ul><li>(A) NAME/KEY:</li><li>(B) LOCATION:</li><li>(D) OTHER INFORMATION: 12R/M13R primer</li></ul>	
72 13 14	30	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:79:	
i Ž		CAGGAAACAG CTATGACCGG ATCCACCTGA GGTCAGAATA	40
		(2) INFORMATION FOR SEQ ID NO:80:	
on Same Vente grade	35	<ul><li>(i) SEQUENCE CHARACTERISTICS:</li><li>(A) LENGTH: 21 base pairs</li><li>(B) TYPE: nucleic acid</li><li>(C) STRANDEDNESS: single</li></ul>	
	40	(D) TOPOLOGY: linear	
		<ul><li>(A) NAME/KEY:</li><li>(B) LOCATION:</li><li>(D) OTHER INFORMATION: 13-2F primer</li></ul>	
	45	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:80:	
		TAACATTTAA GCATCCGTTA C	21
	50	(2) INFORMATION FOR SEQ ID NO:81:	
	55	<ul><li>(i) SEQUENCE CHARACTERISTICS:</li><li>(A) LENGTH: 28 base pairs</li><li>(B) TYPE: nucleic acid</li><li>(C) STRANDEDNESS: single</li><li>(D) TOPOLOGY: linear</li></ul>	
	60	<ul><li>(A) NAME/KEY:</li><li>(B) LOCATION:</li><li>(D) OTHER INFORMATION: 13-2R primer</li></ul>	

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:81:

5	<ul><li>(A) NAME/KEY:</li><li>(B) LOCATION:</li><li>(D) OTHER INFORMATION: 15FUT/M13-R primer</li></ul>	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:85:	
10	CAGGAAACAG CTATGACCAC TCTGTCATAA AAGCCATC	38
	(2) INFORMATION FOR SEQ ID NO:86:	
15	<ul><li>(i) SEQUENCE CHARACTERISTICS:</li><li>(A) LENGTH: 24 base pairs</li><li>(B) TYPE: nucleic acid</li><li>(C) STRANDEDNESS: single</li><li>(D) TOPOLOGY: linear</li></ul>	
20	(B) 1010201	
	<ul><li>(A) NAME/KEY:</li><li>(B) LOCATION:</li><li>(D) OTHER INFORMATION: 16AF primer</li></ul>	
25	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:86:	
		24
	TTTGGTTTGT TATAATTGTT TTTA	24
30	(2) INFORMATION FOR SEQ ID NO:87:	
	<ul><li>(i) SEQUENCE CHARACTERISTICS:</li><li>(A) LENGTH: 20 base pairs</li><li>(B) TYPE: nucleic acid</li></ul>	
35	(C) STRANDEDNESS: single (D) TOPOLOGY: linear	
40	<ul><li>(A) NAME/KEY:</li><li>(B) LOCATION:</li><li>(D) OTHER INFORMATION: 16AR primer</li></ul>	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:87:	
45	CCAACTTTTT AGTTCGAGAG	20
	(2) INFORMATION FOR SEQ ID NO:88:	
50	<ul><li>(i) SEQUENCE CHARACTERISTICS:</li><li>(A) LENGTH: 19 base pairs</li><li>(B) TYPE: nucleic acid</li><li>(C) STRANDEDNESS: single</li><li>(D) TOPOLOGY: linear</li></ul>	
55		
	<ul><li>(A) NAME/KEY:</li><li>(B) LOCATION:</li><li>(D) OTHER INFORMATION: 17F primer</li></ul>	
60	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:88:	
	TTCAGTATCA TCCTATGTG	19

(D) TOPOLOGY: linear

		(2) INFORMATION FOR SEQ ID NO:89:	
	5	<ul><li>(i) SEQUENCE CHARACTERISTICS:</li><li>(A) LENGTH: 20 base pairs</li><li>(B) TYPE: nucleic acid</li></ul>	
		<pre>(C) STRANDEDNESS: single (D) TOPOLOGY: linear</pre>	
	10	(b) Torollog1. Timear	
	15	<ul><li>(A) NAME/KEY:</li><li>(B) LOCATION:</li><li>(D) OTHER INFORMATION: 17AR primer</li></ul>	
	15	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:89:	
		AGAAACCTTA ACCCATACTG	20
	20	(2) INFORMATION FOR SEQ ID NO:90:	
or at treast through the file	25	<ul><li>(i) SEQUENCE CHARACTERISTICS:</li><li>(A) LENGTH: 39 base pairs</li><li>(B) TYPE: nucleic acid</li><li>(C) STRANDEDNESS: single</li><li>(D) TOPOLOGY: linear</li></ul>	
	30	<ul><li>(A) NAME/KEY:</li><li>(B) LOCATION:</li><li>(D) OTHER INFORMATION: 18FUT/M13-AF primer</li></ul>	
		(xi) SEQUENCE DESCRIPTION: SEQ ID NO:90:	
	35	TGTAAAACGA CGGCCAGTGA ATTCTAGAGT CACACTTCC	39
		(2) INFORMATION FOR SEQ ID NO:91:	
	40	<ul><li>(i) SEQUENCE CHARACTERISTICS:</li><li>(A) LENGTH: 38 base pairs</li><li>(B) TYPE: nucleic acid</li><li>(C) STRANDEDNESS: single</li><li>(D) TOPOLOGY: linear</li></ul>	
	45		
		<ul><li>(A) NAME/KEY:</li><li>(B) LOCATION:</li><li>(D) OTHER INFORMATION: 18R/M13R primer</li></ul>	
	50	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:91:	
		CAGGAAACAG CTATGACCTT TAACTGAATC AATGACTG	38
		(2) INFORMATION FOR SEQ ID NO:92:	
	55	(i) SEQUENCE CHARACTERISTICS:  (A) LENGTH: 41 base pairs  (B) TYPE: nucleic acid	
	60	<ul><li>(C) STRANDEDNESS: single</li><li>(D) TOPOLOGY: linear</li></ul>	

5	<ul><li>(A) NAME/KEY:</li><li>(B) LOCATION:</li><li>(D) OTHER INFORMATION: 19F/M13F primer</li></ul>	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:92:	
	TGTAAAACGA CGGCCAGTAA GTGAATATTT TTAAGGCAGT T	41
10	(2) INFORMATION FOR SEQ ID NO:93:	
15	<ul><li>(i) SEQUENCE CHARACTERISTICS:</li><li>(A) LENGTH: 39 base pairs</li><li>(B) TYPE: nucleic acid</li><li>(C) STRANDEDNESS: single</li><li>(D) TOPOLOGY: linear</li></ul>	
20	<ul><li>(A) NAME/KEY:</li><li>(B) LOCATION:</li><li>(D) OTHER INFORMATION: 19FUT/M13-R primer</li></ul>	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:93:	
25	CAGGAAACAG CTATGACCAA GAGACCGAAA CTCCATCTC	39
	(2) INFORMATION FOR SEQ ID NO:94:	
30	<ul><li>(i) SEQUENCE CHARACTERISTICS:</li><li>(A) LENGTH: 38 base pairs</li><li>(B) TYPE: nucleic acid</li><li>(C) STRANDEDNESS: single</li><li>(D) TOPOLOGY: linear</li></ul>	
35		
	<ul><li>(A) NAME/KEY:</li><li>(B) LOCATION:</li><li>(D) OTHER INFORMATION: 20F/M13F primer</li></ul>	
40	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:94:	
	TGTAAAACGA CGGCCAGTCA CTGTGCCTGG CCTGATAC	38
45	(2) INFORMATION FOR SEQ ID NO:95:	
43	<ul><li>(i) SEQUENCE CHARACTERISTICS:</li><li>(A) LENGTH: 39 base pairs</li><li>(B) TYPE: nucleic acid</li></ul>	
50	<ul><li>(C) STRANDEDNESS: single</li><li>(D) TOPOLOGY: linear</li></ul>	
55	<ul><li>(A) NAME/KEY:</li><li>(B) LOCATION:</li><li>(D) OTHER INFORMATION: 20R/M13R primer</li></ul>	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:95:	
60	CAGGAAACAG CTATGACCAT GTTAAATTCA AAGTCTCTA	39
30	(2) INFORMATION FOR SEC ID NO:96:	

	5	<ul><li>(i) SEQUENCE CHARACTERISTICS:</li><li>(A) LENGTH: 39 base pairs</li><li>(B) TYPE: nucleic acid</li><li>(C) STRANDEDNESS: single</li><li>(D) TOPOLOGY: linear</li></ul>	
	10	<ul><li>(A) NAME/KEY:</li><li>(B) LOCATION:</li><li>(D) OTHER INFORMATION: 21F/M13F primer</li></ul>	
		(xi) SEQUENCE DESCRIPTION: SEQ ID NO:96:	
	15	TGTAAAACGA CGGCCAGTGG GTGTTTTATG CTTGGTTCT	39
		(2) INFORMATION FOR SEQ ID NO:97:	
=	20	<ul><li>(i) SEQUENCE CHARACTERISTICS:</li><li>(A) LENGTH: 40 base pairs</li><li>(B) TYPE: nucleic acid</li><li>(C) STRANDEDNESS: single</li><li>(D) TOPOLOGY: linear</li></ul>	
	25	<ul><li>(A) NAME/KEY:</li><li>(B) LOCATION:</li><li>(D) OTHER INFORMATION: 21R/M13R primer</li></ul>	
:	30	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:97:	
	30	CAGGAAACAG CTATGACCCA TTTCAACATA TTCCTTCCTG	40
	35	(2) INFORMATION FOR SEQ ID NO:98:  (i) SEQUENCE CHARACTERISTICS:  (A) LENGTH: 19 base pairs	
	40	<ul><li>(B) TYPE: nucleic acid</li><li>(C) STRANDEDNESS: single</li><li>(D) TOPOLOGY: linear</li></ul>	
	45	<ul><li>(A) NAME/KEY:</li><li>(B) LOCATION:</li><li>(D) OTHER INFORMATION: 22F-1A primer</li></ul>	
		(xi) SEQUENCE DESCRIPTION: SEQ ID NO:98:	19
	50	AACCACACCC TTAAGATGA	19
	55	<ul><li>(2) INFORMATION FOR SEQ ID NO:99:</li><li>(i) SEQUENCE CHARACTERISTICS:</li><li>(A) LENGTH: 20 base pairs</li><li>(B) TYPE: nucleic acid</li></ul>	
		(C) STRANDEDNESS: single (D) TOPOLOGY: linear	
	60	<ul><li>(A) NAME/KEY:</li><li>(B) LOCATION:</li><li>(D) OTHER INFORMATION: 22R-1A primer</li></ul>	

	(X1) SEQUENCE DESCRIPTION. SEQ ID NO. 33.	
5	GCATTAGTAG TGGATTTTGC	20
	(2) INFORMATION FOR SEQ ID NO:100:	
10	<ul><li>(i) SEQUENCE CHARACTERISTICS:</li><li>(A) LENGTH: 16 base pairs</li><li>(B) TYPE: nucleic acid</li><li>(C) STRANDEDNESS: single</li><li>(D) TOPOLOGY: linear</li></ul>	
15	<ul><li>(A) NAME/KEY:</li><li>(B) LOCATION:</li><li>(D) OTHER INFORMATION: 23FII primer</li></ul>	
20	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:100:	
	TCACTTCCAT TGCATC	16
٥٦	(2) INFORMATION FOR SEQ ID NO:101:	
25	<ul><li>(i) SEQUENCE CHARACTERISTICS:</li><li>(A) LENGTH: 17 base pairs</li><li>(B) TYPE: nucleic acid</li><li>(C) STRANDEDNESS: single</li></ul>	
30	(D) TOPOLOGY: linear	
35	(A) NAME/KEY: (B) LOCATION: (D) OTHER INFORMATION: 23RII primer	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:101:	17
40	TGCCAACTGG TAGCTCC  (2) INFORMATION FOR SEQ ID NO:102:	1,
45	<ul><li>(i) SEQUENCE CHARACTERISTICS:</li><li>(A) LENGTH: 20 base pairs</li><li>(B) TYPE: nucleic acid</li><li>(C) STRANDEDNESS: single</li><li>(D) TOPOLOGY: linear</li></ul>	
50	<ul><li>(A) NAME/KEY:</li><li>(B) LOCATION:</li><li>(D) OTHER INFORMATION: 24 2F primer</li></ul>	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:102:	
55	TACAGTTAGC AGCGACAAAA	20
	(2) INFORMATION FOR SEQ ID NO:103:	
60	<ul><li>(i) SEQUENCE CHARACTERISTICS:</li><li>(A) LENGTH: 38 base pairs</li><li>(B) TYPE: nucleic acid</li></ul>	

	<pre>(C) STRANDEDNESS: single (D) TOPOLOGY: linear</pre>	
5	<ul><li>(A) NAME/KEY:</li><li>(B) LOCATION:</li><li>(D) OTHER INFORMATION: 24R/M13R primer</li></ul>	
10	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:103:	
	CAGGAAACAG CTATGACCAT TTGCCAACTG GTAGCTCC	38
15	(2) INFORMATION FOR SEQ ID NO:104:	
20	<ul><li>(i) SEQUENCE CHARACTERISTICS:</li><li>(A) LENGTH: 20 base pairs</li><li>(B) TYPE: nucleic acid</li><li>(C) STRANDEDNESS: single</li><li>(D) TOPOLOGY: linear</li></ul>	
25	<ul><li>(A) NAME/KEY:</li><li>(B) LOCATION:</li><li>(D) OTHER INFORMATION: 25F-7/23 primer</li></ul>	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:104:	
30	GCTTTCGCCA AATTCAGCTA	20
	(2) INFORMATION FOR SEQ ID NO:105:	
35	<ul><li>(i) SEQUENCE CHARACTERISTICS:</li><li>(A) LENGTH: 20 base pairs</li><li>(B) TYPE: nucleic acid</li><li>(C) STRANDEDNESS: single</li><li>(D) TOPOLOGY: linear</li></ul>	
40	<ul><li>(A) NAME/KEY:</li><li>(B) LOCATION:</li><li>(D) OTHER INFORMATION: 25R-7/23 primer</li></ul>	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:105:	
45	TACCAAAATG TGTGGTGATG	20
	(2) INFORMATION FOR SEQ ID NO:106:	
50	<ul><li>(i) SEQUENCE CHARACTERISTICS:</li><li>(A) LENGTH: 20 base pairs</li><li>(B) TYPE: nucleic acid</li><li>(C) STRANDEDNESS: single</li><li>(D) TOPOLOGY: linear</li></ul>	
55		
60	(A) NAME/KEY: (B) LOCATION: (D) OTHER INFORMATION: 26-2F primer	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:106:	

AATCACTGAT ACTGGTTTTG

20

(D) TOPOLOGY: linear

5	<ul><li>(A) NAME/KEY:</li><li>(B) LOCATION:</li><li>(D) OTHER INFORMATION: 27BF/M13F primer</li></ul>	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:110:	
1.0	TGTAAAACGA CGGCCAGTGA ATTCTCCTCA GATGACTCCA	40
10	(2) INFORMATION FOR SEQ ID NO:111:	
15	<ul><li>(i) SEQUENCE CHARACTERISTICS:</li><li>(A) LENGTH: 38 base pairs</li><li>(B) TYPE: nucleic acid</li><li>(C) STRANDEDNESS: single</li><li>(D) TOPOLOGY: linear</li></ul>	
20	(A) NAME/KEY:	
	<ul><li>(B) LOCATION:</li><li>(D) OTHER INFORMATION: 27BR/M13R primer</li></ul>	
٥٦	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:111:	
25	CAGGAAACAG CTATGACCTC TTTGCTCATT GTGCAACA	38

20

25

#### WE CLAIM:

5 1. A genomic DNA containing a BRCA2 gene,

wherein the first twelve nucleotides beginning exon 5 are 5'-

TCCTGTTGTTCT-3' as set forth in SEQ. ID. NO: 1,

wherein nucleotides numbers 5782-5790 are GTTTGTGTT as set forth in SEQ. ID. NO: 4, and

wherein the last 20 nucleotides ending exon 15 are 5'CTGCGTGTTCTCATAAACAG-3' as set forth in SEQ. ID. NO: 2 and the first 20
nucleotides beginning exon 16 are 5'-CTGTATACGTATGGCGTTTC-3' as set forth
in SEQ. ID. NO: 3.

2. The genomic DNA according to claim 1 wherein the coding sequence nucleotides are as follows:

1093 A 1342 A 1593 A 2457 T 2908 G 3199 A 3624 A 4035 T 7470 A 9079 G.

3. The genomic DNA according to claim 1 wherein the coding sequence nucleotides are as follows:

30 1093 A 1342 C 1593 A 2457 T 35 2908 G 3199 A 3624 A 4035 T 7470 A 9079 G.

4. The genomic DNA according to claim 1 wherein the coding sequence nucleotides are as follows:

	1093 A
	1342 C
5	1593 A
	2457 T
	2908 G
	3199 A
	3624 A
10	4035 C
	7470 A
	9079 G.

15 5. The genomic DNA according to claim 1 wherein the coding sequence nucleotides are as follows:

```
1342 A

20 1593 A

2457 C

2908 G

3199 G

3624 G

25 4035 T

7470 G

9079 G.
```

1093 C

6. The genomic DNA according to claim 1 wherein the coding sequence nucleotides are as follows:

```
1342 C
1593 A
35 2457 T
2908 G
3199 A
3624 G
4035 T
40 7470 G
9079 G.
```

1093 A

7. The genomic DNA according to claim 1 wherein the coding sequence nucleotides are as follows:

```
1093 C
1342 C
1593 G
2457 C
50 2908 A
3199 G
```

20

3624 A 4035 T 7470 A 5 9079 A.

8. The genomic DNA according to claim 1 wherein the coding sequence nucleotides are as follows:

2024 C 4553 C 4815 G 5841 T 5972 C.

9. A DNA comprising a BRCA2 coding sequence, wherein nucleotide numbers 643-666 are

CTTAGTGAAAGTCCTGTTGTTCTA and

wherein nucleotides numbers 5782-5790 are GTTTGTGTT.

10. The DNA according to claim 9 wherein the coding sequence nucleotides are as follows:

25 1342 A 1593 A 2457 T 2908 G 3199 A 30 3624 A 4035 T 7470 A 9079 G.

1093 A

11. The DNA according to claim 9 wherein the coding sequence nucleotides are as follows:

1093 A 1342 C 40 1593 A 2457 T 2908 G 3199 A 3624 A 45 4035 T 7470 A 9079 G as set forth in SEQ. ID. NO: 4.

12. The DNA according to claim 9 wherein the coding sequence nucleotides are as follows:

```
1093 A
1342 C
5 1593 A
2457 T
2908 G
3199 A
3624 A
10 4035 C
7470 A
9079 G
as set forth in SEQ. ID. NO: 6.
```

13. The DNA according to claim 9 wherein the coding sequence nucleotides are as follows:

```
1342 A
20 1593 A
2457 C
2908 G
3199 G
3624 G
25 4035 T
7470 G
9079 G
as set forth in SEQ. ID. NO: 8.
```

1093 C

30 14. The DNA according to claim 9 wherein the coding sequence nucleotides are as follows:

```
1342 C

1593 A

2457 T

2908 G

3199 A

3624 G

40 4035 T

7470 G

9079 G

as set forth in SEQ. ID. NO: 10.
```

1093 A

15. The DNA according to claim 9 wherein the coding sequence nucleotides are as follows:

```
1093 C
1342 C
50 1593 G
2457 C
```

25

```
2908 A
         3199 G
          3624 A
         4035 T
5
          7470 A
          9079 A
    as set forth in SEQ. ID. NO: 12.
```

The DNA according to claim 9 wherein the coding sequence nucleotides are 16. 10 as follows:

2024 C 4553 C 15 4815 G 5841 T 5972 C.

289

A BRCA2 protein having the following amino acids at the following peptide 17. numbers:

asparagine 372 histidine 894 valine 991 asparagine 1852 valine 1853 cysteine 1854 valine 2951 alanine

as set forth in SEQ. ID. NO: 5. 30

The BRCA2 protein having the following amino acids at the following peptide 18. numbers:

289 asparagine 35 372 asparagine 599 serine 894 valine 991 asparagine 40 2951 alanine.

> The BRCA2 protein having the following amino acids at the following peptide 19. numbers:

289 histidine 45 histidine 372 valine 894 991 asparatic acid 2951 alanine

as set forth in SEQ. ID. NO: 9. 50

- 20. The BRCA2 protein having the following amino acids at the following peptide numbers:
- 5 289 histidine
  - 372 asparagine
  - 894 isoleucine
  - 991 aspartic acid
  - 2951 threonine
- as set forth in SEQ. ID. NO: 13.
  - 21. The BRCA2 protein according to claims 17-20 having the following amino acids at the following peptide numbers:
- 15 **599 serine** 
  - 1442 serine
  - 1915 threonine.
  - 22. A haplotype of BRCA2 coding sequence (BRCA2° ) as set forth in SEQ. ID.
- NO: 4 or a sequence complementary thereto.
  - 23. A BRCA2 protein comprising an amino acid sequence derived from BRCA2<sup>oml</sup> as set forth in SEQ. ID. NO: 5.
- 25 24. A haplotype of BRCA2 coding sequence (BRCA2<sup>om/2</sup>) as set forth in SEQ. ID. NO: 6 or a sequence complementary thereto.
  - 25. A BRCA2 protein comprising an amino acid sequence derived from BRCA2<sup>omi</sup>
    <sup>2</sup> as set forth in SEQ. ID. NO: 7.
  - 26. A haplotype of BRCA2 coding sequence (BRCA2° ) as set forth in SEQ. ID. NO: 8 or a sequence complementary thereto.
- 27. A BRCA2 protein comprising an amino acid sequence derived from BRCA2<sup>oml</sup> 35 3 as set forth in SEQ. ID. NO: 9.

15

20

25

- 28. A haplotype of BRCA2 coding sequence (BRCA2<sup>om 4</sup>) as set forth in SEQ. ID. NO: 10 or a sequence complementary thereto.
- <sup>4</sup> as set forth in SEQ. ID. NO: 11.
  - 30. A haplotype of BRCA2 coding sequence (BRCA2<sup>om, 5</sup>) as set forth in SEQ. ID. NO: 12 or a sequence complementary thereto.
  - 31. A BRCA2 protein comprising an amino acid sequence derived from BRCA2<sup>omi</sup> s set forth in SEQ. ID. NO: 13.
  - 32. A method of identifying individuals having a BRCA2 gene with a BRCA2 coding sequence not associated with disease, comprising:
    - (a) amplifying a DNA or a fragment thereof of an individual's BRCA2 coding sequence;
    - (b) sequencing said amplified DNA fragment;
    - (c) if necessary, repeating steps (a) and (b) until said individual's BRCA2 coding sequence is sufficiently sequenced to determine whether a mutation is present;
    - (d) comparing the sequence of said amplified DNA fragment to a
       BRCA2<sup>(omi)</sup> DNA sequence selecting from the group consisting of SEQ.
       ID. NO: 4, SEQ. ID. NO: 6, SEQ. ID. NO: 8, SEQ. ID. NO: 10, SEQ. ID.
       NO: 12, and their respective complementary sequences;
    - (e) determining the presence of absence of each of the following polymorphic variations in said individual's BRCA2 coding sequence:
      - (i) AAT and CAT at position 1093,
      - (ii) CAT and AAT at position 1342,
      - (iii) TCA and TCG at position 1593,

10

15

25

- (iv) CAT and CAC at position 2457,
- (v) GTA and ATA at position 2908,
- (vi) AAC and GAC at position 3199,
- (vii) AAA and AAG at position 3624,
- (viii) GTT and GTC at position 4035,
- (ix) TCA and TCG at position 7470, and
- (x) GCC and ACC at position 9079; and
- (f) determining any sequence differences between said individual's BRCA2 coding sequences and a BRCA2<sup>(orm)</sup> DNA sequence selected from the group consisting of SEQ. ID. NO: 4, SEQ. ID. NO: 6, SEQ. ID. NO: 8, SEQ. ID. NO: 10, SEQ. ID. NO: 12, and their respective complementary sequences, wherein the presence of said polymorphic variations and the absence of a variation outside of positions 1093, 1342, 1593, 2457, 2908, 3199, 3624, 4035, 7470, and 9079 is correlated with an absence of increased genetic susceptibility to breast or ovarian cancer resulting from a BRCA2 mutation in the BRCA2 coding sequence.
- 33. A method of identifying individuals having a BRCA2 gene with a BRCA2 coding sequence not associated with disease, comprising:
  - (a) amplifying a DNA or a fragment thereof of an individual's BRCA2 coding sequence;
  - (b) sequencing said amplified DNA fragment;
  - (c) if necessary, repeating steps (a) and (b) until said individual's BRCA2 coding sequence is sufficiently sequenced to determine whether a mutation is present;
  - (d) comparing the sequence of said amplified DNA fragment to a BRCA2<sup>(orni)</sup> DNA sequence selecting from the group consisting of SEQ.
     ID. NO: 4, SEQ. ID. NO: 6, SEQ. ID. NO: 8, SEQ. ID. NO: 10, SEQ. ID. NO: 12, and their respective complementary sequences;

10

15

20

- (e) determining the presence of absence of each of the following polymorphic variations in said individual's BRCA2 coding sequence:
  - (i) AAT and CAT at position 1093,
  - (ii) CAT and AAT at position 1342,
  - (iii) TCA and TCG at position 1593,
  - (iv) CAT and CAC at position 2457,
  - (v) GTA and ATA at position 2908,
  - (vi) AAC and GAC at position 3199,
  - (vii) AAA and AAG at position 3624,
  - (viii) GTT and GTC at position 4035,
  - (ix) TCA and TCG at position 7470, and
  - (x) GCC and ACC at position 9079; and
- (f) determining any sequence differences between said individual's BRCA2 coding sequences and a BRCA2<sup>(orm)</sup> DNA sequence selected from the group consisting of SEQ. ID. NO: 4, SEQ. ID. NO: 6, SEQ. ID. NO: 8, SEQ. ID. NO: 10, SEQ. ID. NO: 12, and their respective complementary sequences, wherein the presence of said polymorphic variations and the absence of a variation outside of positions 1093, 1342, 1593, 2457, 2908, 3199, 3624, 4035, 7470, and 9079 is correlated with an absence of increased genetic susceptibility to breast or ovarian cancer resulting from a BRCA2 mutation in the BRCA2 coding sequence; wherein, codon variations occur at the following frequencies, respectively, in a Caucasian population of individuals free of disease:
  - (i) at position 1093, AAT and CAT occur at frequencies from about 75-85%, and from about 15-25%, respectively,
  - (ii) at position 1342, CAT and AAT occur at frequencies from about 35-45%, and from about 55-65%, respectively.
  - (iii) at position 1593, TCA and TCG occur at frequencies from about 85-95%, and from about 5-15%, respectively,

25

\_\_\_

25

(iv) at position 2457, CAT and CAC occur at frequencies from about 75-85%, and from about 15-25%, respectively, at position 2908, GTA and ATA occur at frequencies from (v) 5 about 85-95%, and from about 5-15%, respectively. (vi) at position 3199, AAC and GAC occur at frequencies from about 75-85%, and from about 15-25%. respectively. (vii) at position 3624, AAA and AAG occur at frequencies 10 from about 75-85%, and from about 15-25%. respectively, (viii) at position 4035, GTT and GTC occur at frequencies from about 85-95%, and from about 5-15%, respectively, (ix) at position 7470, TCA and TCG occur at frequencies from

- 34. A method of detecting an increased genetic susceptibility to breast and ovarian cancer in an individual resulting from the presence of a mutation in the BRCA2 coding sequence, comprising:
  - (a) amplifying a DNA or a fragment thereof of an individual's BRCA2 coding sequence;

about 75-85%, and from about 15-25%, respectively, and

from about 85-95%, and from about 5-15%, respectively.

at position 9079, GCC and ACC occur at frequencies

(b) sequencing said amplified DNA fragment;

(x)

- (c) if necessary, repeating steps (a) and (b) until said individual's BRCA2 coding sequence is sufficiently sequenced to determine whether a mutation is present;
- (d) comparing the sequence of said amplified DNA fragment to a BRCA2<sup>(omi)</sup> DNA sequence selected from the group consisting of SEQ.

10

15

20

- ID. NO: 4, SEQ. ID. NO: 6, SEQ. ID. NO: 8, SEQ. ID. NO: 10, SEQ. ID. NO: 12, and their respective complementary sequences;
- (e) determining any sequence differences between said individual's BRCA2 coding sequences and a BRCA2<sup>(omi)</sup> DNA sequence selected from the group consisting of SEQ. ID. NO: 4, SEQ. ID. NO: 6, SEQ. ID. NO: 8, SEQ. ID. NO: 10, SEQ. ID. NO: 12, and their respective complementary sequences in order to determine the presence or absence of base changes in said individual's BRCA2 coding sequence wherein a base change which is not any one of the following:
  - (i) AAT and CAT at position 1093.
  - (ii) CAT and AAT at position 1342,
  - (iii) TCA and TCG at position 1593,
  - (iv) CAT and CAC at position 2457,
  - (v) GTA and ATA at position 2908,
  - (vi) AAC and GAC at position 3199,
  - (vii) AAA and AAG at position 3624,
  - (viii) GTT and GTC at position 4035.
  - (ix) TCA and TCG at position 7470, and
  - (x) GCC and ACC at position 9079, is correlated with the potential of increased genetic susceptibility to breast or ovarian cancer resulting from a BRCA2 mutation in the BRCA2 coding sequence.
- 35. A method of detecting an increased genetic susceptibility to breast and ovarian cancer in an individual resulting from the presence of a mutation in the BRCA2 coding sequence, comprising:
  - (a) amplifying a DNA or a fragment thereof of an individual's BRCA2 coding sequence;
  - (b) sequencing said amplified DNA fragment;

- (c) if necessary, repeating steps (a) and (b) until said individual's BRCA2 coding sequence is sufficiently sequenced to determine whether a mutation is present;
- (d) comparing the sequence of said amplified DNA fragment to a BRCA2<sup>(orni)</sup> DNA sequence selected from the group consisting of: SEQ.
   ID. NO: 4, SEQ. ID. NO: 6, SEQ. ID. NO: 8, SEQ. ID. NO: 10, SEQ. ID. NO: 12, and their respective complementary sequences;
- (e) determining any sequence differences between said individual's BRCA2 coding sequences and a BRCA2<sup>(orni)</sup> DNA sequence selected from the group consisting of: SEQ. ID. NO: 4, SEQ. ID. NO: 6, SEQ. ID. NO: 8, SEQ. ID. NO: 10, SEQ. ID. NO: 12, and their respective complementary sequences in order to determine the presence or absence of base changes in said individual's BRCA2 coding sequence wherein a base change which is not any one of the following:
  - (i) AAT and CAT at position 1093,
  - (ii) CAT and AAT at position 1342.
  - (iii) TCA and TCG at position 1593,
  - (iv) CAT and CAC at position 2457,
  - (v) GTA and ATA at position 2908,
  - (vi) AAC and GAC at position 3199,
  - (vii) AAA and AAG at position 3624,
  - (viii) GTT and GTC at position 4035,
  - (ix) TCA and TCG at position 7470, and
  - (x) GCC and ACC at position 9079, is correlated with the potential of increased genetic susceptibility to breast or ovarian cancer resulting from a BRCA2 mutation in the BRCA2 coding sequence, wherein, codon variations occur at the following frequencies, respectively, in a Caucasian population of individuals free of disease:

15

20

- (i) at position 1093, AAT and CAT occur at frequencies from about 75-85%, and from about 15-25%, respectively. (ii) at position 1342, CAT and AAT occur at frequencies from 5 about 35-45%, and from about 55-65%, respectively, (iii) at position 1593, TCA and TCG occur at frequencies from about 85-95%, and from about 5-15%, respectively. (iv) at position 2457, CAT and CAC occur at frequencies from about 75-85%, and from about 15-25%, respectively, 10 (v) at position 2908, GTA and ATA occur at frequencies from about 85-95%, and from about 5-15%, respectively, (vi) at position 3199, AAC and GAC occur at frequencies from about 75-85%, and from about 15-25%, respectively, 15 (vii) at position 3624, AAA and AAG occur at frequencies from about 75-85%, and from about 15-25%, respectively, (viii) at position 4035, GTT and GTC occur at frequencies from about 85-95%, and from about 5-15%, respectively. at position 7470, TCA and TCG occur at frequencies from 20 (ix) about 75-85%, and from about 15-25%, respectively, and at position 9079, GCC and ACC occur at frequencies (x)from about 85-95%, and from about 5-15%, respectively.
- 36. A method according to any of the claims 32-35 wherein the said amplifying is performed by annealing at least one oligonucleotide primer to said DNA fragment and extending the oligonucleotide primer by an agent for polymerization.

- 37. A method according to claim 36 wherein said oligonucleotide primer is directly or indirectly labeled with a radioactive label, a fluorescent label, a bioluminescent label, a chemiluminescent label, a metal chelator, or an enzyme label.
- 38. A BRCA2 coding sequence according to claims 32, wherein the codon pairs occur at the following frequencies:
  - (i) at position 1093, AAT and CAT occur at frequencies from about 75-85%, and from about 15-25%, respectively,
  - (ii) at position 1342, CAT and AAT occur at frequencies from about 35-45%, and from about 55-65%, respectively.
  - (iii) at position 1593, TCA and TCG occur at frequencies from about 85-95%, and from about 5-15%, respectively,
  - (iv) at position 2457, CAT and CAC occur at frequencies from about 75-85%, and from about 15-25%, respectively,
  - (v) at position 2908, GTA and ATA occur at frequencies from about 85-95%, and from about 5-15%, respectively,
  - (vi) at position 3199, AAC and GAC occur at frequencies from about 75-85%, and from about 15-25%, respectively,
  - (vii) at position 3624, AAA and AAG occur at frequencies from about 75-85%, and from about 15-25%, respectively,
  - (viii) at position 4035, GTT and GTC occur at frequencies from about 85-95%, and from about 5-15%, respectively,
  - (ix) at position 7470, TCA and TCG occur at frequencies from about 75-85%, and from about 15-25%, respectively, and
  - (x) at position 9079, GCC and ACC occur at frequencies from about 85-95%, and from about 5-15%, respectively.

15

20

10

15

20

30

- 39. An oligonucleotide primer capable of hybridizing to a sample of BRCA2 gene, or its respective complementary sequences selected from the group consisting of SEQ. ID. NO: 14, 19, 22, 23, 25, 26, 29-76, 83, 85-88, 90, 91, 97, 98, 101, and 104-107.
- 40. A chip array having "n" elements for performing allele specific sequence-based techniques comprising a solid phase chip and oligonucleotides having "n" different nucleotide sequences,

wherein "n" is an interger greater than or equal to ten,

wherein said oligonucleotides are bound to said solid phase chip in a manner which permits said oligonucleotides to effectively hybridize to complementary oligonucleotides or polynucleotides,

wherein oligonucleotides having different nucleotide sequence are bound to said solid phase chip at different locations so that a particular location on said solid phase chip exclusively binds oligonucleotides having a specific nucleotide sequence, and

wherein at least ten oligonucleotides are capable of specifically hybridizing to the BRCA2 DNA having the sequence as set forth in SEQ. ID. NO: 4, SEQ. ID. NO: 6, SEQ. ID. NO: 10, SEQ. ID. NO: 12 or their respective complementary sequences, at least one oligonucleotide being capable of specifically hybridizing at each of the nucleotide positions 1093, 1342, 1593, 2457, 2908, 3199, 3624, 4035, 7470, 9079, or complementary thereto.

- 25 41. A method of performing gene therapy on a patient, comprising:
  - a) contacting cancer cells *in vivo* with an effective amount of a vector comprising DNA containing at least a portion of BRCA2 sequence selected from the group consisting of SEQ. ID. NO: 4, SEQ. ID. NO: 6, SEQ. ID. NO: 8, SEQ. ID. NO: 10, SEQ. ID. NO: 12, or their respective complementary sequences
    - b) allowing the vector to enter the cancer cells, and
    - c) measuring a reduction in tumor growth.
  - 42. The method according to claim 41 wherein said cancer cells have a mutation in the BRCA2 gene.

25

5

10

- 43. The method according to claim 41 wherein said patient has a mutation in the BRCA2 gene of non-cancer cells.
- 44. A method of performing gene therapy on a patient or a sample, comprising:
- a) contacting cells *in vivo* or *in vitro* with an effective amount of a vector comprising DNA containing at least a portion of BRCA2 sequence selected from the group consisting of SEQ. ID. NO: 4, SEQ. ID. NO: 6, SEQ. ID. NO: 8, SEQ. ID. NO:
- 10, SEQ. ID. NO: 12, or their respective complementary sequences, and
  - b) allowing the vector to enter the cells,

wherein said patient has a reduced susceptibility for developing a cancer associated with a mutation in the BRCA2 gene.

- 15 45. A method according to claim 44 wherein said cells include healthy breast, ovarian or pancreatic tissues.
  - 46. A method according to claim 44 wherein a patient has an inherited mutation in the BRCA2 gene.

47. A method of treating a patient suspected of having a tumor, comprising:

- a) administering to a patient an effective amount of BRCA2 tumor growth inhibitor having an amino acid sequence selected from the group consisting of SEQ. ID. NO: 5, SEQ. ID. NO: 7, SEQ. ID. NO: 9, SEQ. ID. NO: 11, SEQ. ID. NO: 13, any fragments thereto, and any functional equivalent thereof;
  - b) allowing the patient's cells to take up the protein, and
  - c) measuring a reduction in tumor growth.
- 48. The method according to claim 47 wherein said tumor is a breast cancer, an ovarian cancer or a pancreatic cancer.
  - 49. The method according to claim 47 wherein said patient has an inherited mutation in the BRCA2 gene.

15

- 50. A method of preventing the formation or growth of a tumor, comprising:
- a) adminstering to a patient an effective amount of BRCA2 tumor growth inhibiting protein having an amino acid sequence selected from the group consisting of SEQ. ID. NO: 5, SEQ. ID. NO: 7, SEQ. ID. NO: 9, SEQ. ID. NO: 11, SEQ. ID. NO: 13, any fragments thereto, and any functional equivalent thereof; and
  - b) allowing the patient cells to take up the protein.
- 51. The method according to claim 31 wherein the protein is administered parenternally, by buccal adsorption or inhalation.
  - 52. A cloning vector comprising:
  - (a) a DNA sequence as set forth in SEQ. ID. NO: 4, SEQ. ID. NO: 6, SEQ. ID. NO: 8, SEQ. ID. NO: 10, SEQ. ID. NO: 12, or any fragments thereof; and
  - (b) one or more suitable regulatory sequences to induce replication and/or integration in a host cell.
  - 53. An expression vector comprising a DNA sequence as set forth in SEQ. ID. NO: 4, SEQ. ID. NO: 6, SEQ. ID. NO: 8, SEQ. ID. NO: 10, SEQ. ID. NO: 12, or any fragments thereof operatively linked to one or more promoter sequences capable of directing expression of said sequence in a host cell.
  - 54. A host cell transformed with the vector according to claim 52 or 53.
- A BRCA2 polypeptide which is selected from the group consisting of:

  (a) a fragment of BRCA2 protein sequence as set forth in SEQ. ID. NO: 5,

  SEQ. ID. NO: 7, SEQ. ID. NO: 9, SEQ. ID. NO: 11, or SEQ. ID. NO:13;

  (b) an amino acid sequence which is substantially homologous to the BRCA2 protein sequence as set forth in SEQ. ID. NO: 5, SEQ. ID. NO: 7, SEQ. ID.

  NO: 9, SEQ. ID. NO: 11, or SEQ. ID. NO: 13;

  (c) a molecule which has similar function to the BRCA2 protein; and

  (d) a fusion protein of (a), (b), or (c).

- 56. An anti-BRCA2 antibody wherein a molecule according to claims 17-21, 23, 25, 27, 29, 31, or 55 is used as an immunogen.
- 5 57. A diagnostic reagent comprising a molecule selected from the group consisting of:
  - (a) a DNA sequence as set forth in SEQ. ID. NO: 4, SEQ. ID. NO: 6, SEQ. ID. NO:
  - 8, SEQ. ID. NO: 10, SEQ. ID. NO: 12, or their complementary sequences;
  - (b) a nucleic acid fragment of (a) comprising at least 10 nucleotide in length;
- 10 (c) a sequence which hybridizes to (a) or (b);
  - (d) a polypeptide according to claim 17-21, 23, 25, 27, 29, 31, or 55; and
  - (e) an antibody which specifically binds to the polypeptide of (d).
- 58. A pharmaceutical composition comprising a molecule according to any one of the claims 17-21, 23, 25, 27, 29, 31, 55 in a pharmaceutically acceptable carrier.
  - 59. A pharmaceutical composition comprising a molecule according claim 56 in a pharmaceutically acceptable carrier.
- 20 60. A pharmaceutical composition comprising a molecule according to claim 57 in a pharmaceutically acceptable carrier.

10

# ABSTRACT OF THE DISCLOSURE

Five novel DNA and protein sequences have been determined for the BRCA2 gene, as have been ten polymorphic sites and their rates of occurrence in the normal alleles of BRCA2. The sequences BRCA2<sup>(orn; 1-5)</sup> and the ten polymorphic sites will provide greater accuracy and reliability for genetic testing. One skilled in the art will be better able to avoid misinterpretations of changes in the gene and/or protein sequence, determine the presence of a normal sequence, and of mutations of BRCA2. This invention is also related to a method of performing gene therapy with BRCA2<sup>(orn; 1-5)</sup> coding sequences or fragments thereof. This invention is further related to protein therapy with BRCA2<sup>(orn; 1-5)</sup> proteins or their functional equivalents.

# Figure 1A

Exon 2

taagtgcattttggtcttctgttttgcagACTTATTTACCAAGCATTGGAGGAATATCGTAGGTAAAA<u>ATĞ</u>ČCTATTĞGATCČAAĞGAGAGCCAACATTTTTTGAAATTTTTAAGACACGCTGC AACAAAGCAGgtattgacaaattttatataac

Exon 3

gggattttttttttaaatagATTTAGGACCAATAAGTCTTAATTGGTTTGAAGAACTTTCTTCAG ĂĂĞCTCCACCCTATAATTCTGAACCTGCAGAAGAATCTGAACATAAAAACAACAATT ACGAACCAAACCTATTTAAAACTCCACAAAGGAAACCATCTTATAATCAGCTGGCTT CAACTCCAATAATATTCAAAGAGCAAGGGCTGACTCTGCCGCTGTACCAATCTCCT GTAAAAGAATTAGATAAATTCAAATTAGACTTAGgtaagtaatgcaatatggtagactgggg

Exon 4

 $t cact ga att att gtact gttt cag {\tt GAAGGAATGTTCCCAATAGTAGACATAAAAGTCTTCGCACA}$ GTGĂAAACTAAĂATGĞATCAAGCAGATGATGTTTCCTGTCCACTTCTAAATTCTTGT CTTAGTGAAAGgtatgatgaagctattatattaaaa

Exon 5

agggatttgctttgtTTTATTTTAGTCCTGTTGTTCTACAATGTACACATGTAACACCACAAA GAGATAAGTCAGgtatgattaaaaacaatgctttttattctt

Exon 6

ttaacaattttcccctttttttacccccagTGGTATGTGGGAGTTTGTTTCATACACCAAAGTTTGTG **AAGgtaaatatt** 

Exon 7

TCTGAAAGTCTAGGAGCTGAGGTGGATCCTGATATGTCTTGGTCAAGTTCTTTAGC TACACCACCCACCCTTAGTTCTACTGTGCTCATAGgtaataata

Exon 8

ttttatcttacagTCAGAAATGAAGAAGCATCTGAAACTGTATTTCCTCATGATACTACTGC Tgtaagtaaatatgacattgattagact

Exon 9

taaactataatttttgcagAATGTGAAAAGCTATTTTTCCAATCATGATGAAAGTCTGAAGAAA AATGATAGATTTATCGCTTCTGTGACAGACAGTGAAAACACAAATCAAAGAGAAGC TGCAAGTCATGgtaagtcctct

Exon 10

tta at gtg ctt ctg tttt at a cttta a cag GATTTGGAAAAACATCAGGGAATTCATTTAAAGTAAATAGCTĞCAAĂGACCACATTĞGAAAGTCAATGCCAAATGTCCTAGAAGATĢAAGTATAT GAAACAGTTGTAGATACCTCTGAAGAAGATAGTTTTTCATTATGTTTTTCTAAATGTA GAACAAAAATCTACAAAAAGTAAGAACTAGCAAGACTAGGAAAAAAATTTTCCATG TTGTATCTGAAGTGGAACCAAATGATACTGATCCATTAGATTCAAATGTAGCAAATC

# Figure 1B

# Exon 11

tttgtgtttttatgtttagGTTTATTGCATTCTTCTGTGAAAAGAAGCTGTTCACAGAATGATTCT GĂĂGAAČCAĂCTTTGTCCTTAACTAGCTCTTTTGGGACAATTCTGAGGAAATGTTCT AGAAATGAAACATGTTCTAATAATACAGTAATCTCTCAGGATCTTGATTATAAAGAA GCAAAATGTAATAAGGAAAAACTACAGTTATTTATTACCCCAGAAGCTGATTCTCTG TCATGCCTGCAGGAAGGACAGTGTGAAAATGATCCAAAAAGCAAAAAGTTTCAGA TATAAAAGAAGAGGTCTTGGCTGCAGCATGTCACCCAGTACAACATTCAAAAGTGG AATACAGTGATACTGACTTTCAATCCCAGAAAAGTCTTTTATATGATCATGAAAATG CCAGCACTCTTATTTTAACTCCTACTTCCAAGGATGTTCTGTCAAACCTAGTCATGA TTTCTAGAGGCAAAGAATCATACAAAATGTCAGACAAGCTCAAAGGTAACAATTATG CTTTAAATGAAAATTATAAAAACGTTGAGCTGTTGCCACCTGAAAAATACATGAGAG TAGCATCACCTTCAAGAAAGGTACAATTCAACCAAAACACAAATCTAAGAGTAATCC AAAAAAATCAAGAAGAAACTACTTCAATTTCAAAAATAACTGTCAATCCAGACTCTG AAGAACTTTCTCAGACAATGAGAATAATTTTGTCTTCCAAGTAGCTAATGAAAGGA ATAATCTTGCTTTAGGAAATACTAAGGAACTTCATGAAACAGACTTGACTTGTGTAA ACGAACCCATTTTCAAGAACTCTACCATGGTTTTATATGGAGACACAGGTGATAAAC AAGCAACCCAAGTGTCAATTAAAAAAGATTTGGTTTATGTTCTTGCAGAGGAGAAC AAAAATAGTGTAAAGCAGCATATAAAAATGACTCTAGGTCAAGATTTAAAATCGGAC ATCTCCTTGAATATAGATAAAATACCAGAAAAAAATAATGATTACATGAACAAATGG GCAGGACTCTTAGGTCCAATTTCAAATCACAGTTTTGGAGGTAGCTTCAGAACAGC TTCAAATAAGGAAATCAAGCTCTCTGAACATAACATTAAGAAGAGCAAAATGTTCTT CAAAGATATTGAAGAACAATATCCTACTAGTTTAGCTTGTGTGAAATTGTAAATAC CTTGGCATTAGATAATCAAAAGAAACTGAGCAAGCCTCAGTCAATTAATACTGTATC TGCACATTTACAGAGTAGTGTAGTTGTTCTGATTGTAAAAATAGTCATATAACCCC TCAGATGTTATTTCCAAGCAGGATTTTAATTCAAACCATAATTTAACACCTAGCCAA AAGGCAGAAATTACAGAACTTTCTACTATATTAGAAGAATCAGGAAGTCAGTTTGAA TTTACTCAGTTTAGAAAACCAAGCTACATATTGCAGAAGAGTACATTTGAAGTGCCT GAAAACCAGATGACTATCTTAAAGACCACTTCTGAGGAATGCAGAGATGCTGATCT 

# Figure 1C

AAGGTACAGTTGAAATTAAACGGAAGTTTGCTGGCCTGTTGAAAAATGACTGTAAC AAAAGTGCTTCTGGTTATTTAACAGATGAAAATGAAGTGGGGTTTAGGGGGCTTTTAT TCTGCTCATGGCACAAAACTGAATGTTTCTACTGAAGCTCTGCAAAAAGCTGTGAA ACTGTTTAGTGATATTGAGAATATTAGTGAGGAAACTTCTGCAGAGGTACATCCAAT AAGTTTATCTTCAAGTAAATGTCATGATTCTGTTGTTTCAATGTTTAAGATAGAAAAT CATAATGATAAACTGTAAGTGAAAAAAATAATAAATGCCAACTGATATTACAAAATA ATATTGAAATGACTACTGGCACTTTTGTTGAAGAAATTACTGAAAATTACAAGAGAA ATACTGAAAATGAAGATAACAAATATACTGCTGCCAGTAGAAATTCTCATAACTTAG AATTTGATGGCAGTGATTCAAGTAAAAATGATACTGTTTGTATTCATAAAGATGAAA CGGACTTGCTATTTACTGATCAGCACAACATATGTCTTAAATTATCTGGCCAGTTTA TGAAGGAGGAAACACTCAGATTAAAGAAGATTTGTCAGATTTAACTTTTTTGGAAG TTGCGAAGCTCAAGAAGCATGTCATGGTAATACTTCAAATAAAGAACAGTTAACT GCTACTAAAACGGAGCAAAATATAAAAGATTTTGAGACTTCTGATACATTTTTTCAG ACTGCAAGTGGGAAAAATATTAGTGTCGCCAAAGAGTCATTTAATAAAATTGTAAAT TTCTTTGATCAGAAACCAGAAGAATTGCATAACTTTTCCTTAAATTCTGAATTACATT CTGACATAAGAAAGAACAAAATGGACATTCTAAGTTATGAGGAAACAGACATAGTT AAACACAAAATACTGAAAGAAGTGTCCCAGTTGGTACTGGAAATCAACTAGTGAC CTTCCAGGGACACCCGAACGTGATGAAAAGATCAAAGAACCTACTCTGTTGGGTT TTCATACAGCTAGCGGGAAAAAAGTTAAAATTGCAAAGGAATCTTTGGACAAAGTG AAAAACCTTTTTGATGAAAAAGAGCAAGGTACTAGTGAAATCACCAGTTTTAGCCAT CAATGGGCAAAGACCCTAAAGTACAGAGAGGCCTGTAAAGACCTTGAATTAGCAT GTGAGACCATTGAGATCACAGCTGCCCCAAAGTGTAAAGAAATGCAGAATTCTCTC AATAATGATAAAAACCTTGTTTCTATTGAGACTGTGGTGCCACCTAAGCTCTTAAGT GATAATTTATGTAGACAAACTGAAAATCTCAAAACATCAAAAAGTATCTTTTTGAAAG TTAAAGTACATGAAAATGTAGAAAAAGAAACAGCAAAAAGTCCTGCAACTTGTTACA CAAATCAGTCCCCTTATTCAGTCATTGAAAATTCAGCCTTAGCTTTTTACACAAGTT GTAGTAGAAAACTTCTGTGAGTCAGACTTCATTACTTGAAGCAAAAAAATGGCTTA GAGAAGGAATATTTGATGGTCAACCAGAAAGAATAAATACTGCAGATTATGTAGGA AATTATTTGTATGAAAATAATTCAAACAGTACTATAGCTGAAAATGACAAAAATCATC TCTCCGAAAAACAAGATACTTATTTAAGTAACAGTAGCATGTCTAACAGCTATTCCT ACCATTCTGATGAGGTATATAATGATTCAGGATATCTCTCAAAAAATAAACTTGATT CTGGTATTGAGCCAGTATTGAAGAATGTTGAAGATCAAAAAAACACTAGTTTTTCCA AAGTAATATCCAATGTAAAAGATGCAAATGCATACCCACAAACTGTAAATGAAGATA TTTGCGTTGAGGAACTTGTGACTAGCTCTTCACCCTGCAAAAATAAAAATGCAGCC ATTAAATTGTCCATATCTAATAGTAATATTTTGAGGTAGGGCCACCTGCATTTAGG ATAGCCAGTGGTAAAATCGTTTGTGTTTCACATGAAACAATTAAAAAAAGTGAAAGAC ATATTTACAGACAGTTTCAGTAAAGTAATTAAGGAAAACAACGAGAATAAATCAAAA ATTTGCCAAACGAAAATTATGGCAGGTTGTTACGAGGCATTGGATGATTCAGAGGA TATTCTTCATAACTCTCTAGATAATGATGAATGTAGCACGCATTCACATAAGGTTTTT GCTGACATTCAGAGTGAAGAATTTTACAACATAACCAAAATATGTCTGGATTGGA GAAAGTTTCTAAAATATCACCTTGTGATGTTAGTTTGGAAACTTCAGATATATGTAAA TGTAGTATAGGGAAGCTTCATAAGTCAGTCTCATCTGCAAATACTTGTGGGATTTTT AGCACAGCAAGTGGAAAATCTGTCCAGGTATCAGATGCTTCATTACAAAACGCAAG ACAAGTGTTTTCTGAAATAGAAGATAGTACCAAGCAAGTCTTTTCCAAAGTATTGTT TAAAAGTAACGAACATTCAGACCAGCTCACAAGAGAAAAATACTGCTATACGTA CTCCAGAACATTTAATATCCCAAAAAGGCTTTTCATATAATGTGGTAAATTCATCTG

# Figure 1D

CTTTCTCTGGATTTAGTACAGCAAGTGGAAAGCAAGTTTCCATTTTAGAAAGTTCCT TACACAAAGTTAAGGGAGTGTTAGAGGAATTTGATTTAATCAGAACTGAGCATAGT CTTCACTATTCACCTACGTCTAGACAAAATGTATCAAAAATACTTCCTCGTGTTGAT AAGAGAAACCCAGAGCACTGTGTAAACTCAGAAATGGAAAAAACCTGCAGTAAAGA ATTTAAATTATCAAATAACTTAAATGTTGAAGGTGGTTCTTCAGAAAATAATCACTCT ATTAAAGTTTCTCCATATCTCTCAATTTCAACAAGACAACAACAGTTGGTATTAG GAACCAAAGTCTCACTTGTTGAGAACATTCATGTTTTGGGAAAAGAACAGGCTTCA CCTAAAAACGTAAAAATGGAAATTGGTAAAACTGAAACTTTTTCTGATGTTCCTGTG AAAACAAATATAGAAGTTTGTTCTACTTACTCCAAAGATTCAGAAAACTACTTTGAAA CAGAAGCAGTAGAAATTGCTAAAGCTTTTATGGAAGATGATGAACTGACAGATTCT AAACTGCCAAGTCATGCCACACATTCTCTTTTTACATGTCCCGAAAATGAGGAAATG GTTTTGTCAAATTCAAGAATTGGAAAAAGAAGAGGGGGGGCCCCTTATCTTAGTGGgt aagtgttcatttttacctttcgtgttgccaatca

# Exon 12

aaaacatatatgaaatatttctttttagGAGAACCCTCAATCAAAAGAAACTTATTAAATGAATTTG ACAGGATAATAGAAAATČAAGAAAAATCCTTAAAGGCTTCAAAAAGCACTCCAGAT Ggtaaaattagctttttatttata

#### Exon 13

aatatgtaatataaaataattgtttcctagGCACAATAAAAGATCGAAGATTGTTTATGCATCATGT TTCŤTTAGAGCCGAŤTACCŤGTGTACCCTTTCGgtaagacatgtttaaatttttctaa

## Exon 14

cccattgcagCACAACTAAGGAACGTCAAGAGATACAGAATCCAAATTTTACCGCACC TGGTČAAĞAATTTCTGTCTAAATCTCATTTGTATGAACATCTGACTTTGGAAAAATCT TCAAGCAATTTAGCAGTTTCAGGACATCCATTTTATCAAGTTTCTGCTACAAGAAAT GAAAAAATGAGACACTTGATTACTACAGGCAGACCAAACCAAAGTCTTTGTTCCACC TTTTAAAACTAAATCaCATTTTCACAGAGTTGAACAGTGTGTTAGGAATATTAACTTG GAGGAAACAGACAAAGCAAAACATTGATGGACATGGCTCTGATGATAAAAA TAAGATTAATGACAATGAGATTCATCAGTTTAACAAAAACAACTCCAATCAAGCAGC AGCTGTAACTTTCACAAAGTGTGAAGAAGAACCTTTAGgtattgtatgacaatttgtgtgatgaatt

#### Exon 15

tttttgctaagtatttattctttgatagATTTAATTACAAGTCTTCAGAATGCCAGAGATATACAGGAT ATĞCGĂATTAAGĂAGĂAACAAAGGCAACGCGTCTTTCCACAGCCAGGCAGTCTGTA TCTTGCAAAAACATCCACTCTGCCTCGAATCTCTCTGAAAGCAGCAGTAGGAGGCC  ${\sf AAGTTCCCTCTGCgtgtccccataaacaggtatgtgt}$ 

#### Exon 16

tttttcttttttgtgtgtgtttattttgtgtag GTGTTCTCATAAACAG CTGTATACGTATGGCGTTTCTAAACATTGCATAAAAATTAACAGCAAAAATGCAGAGTCTTTTCAGTTTCACACTGAAGA TTATTTTGGTAAGGAAAGTTTATGGACTGGAAAAGGAATACAGTTGGCTGATGGTG GATGGCTCATACCCTCCAATGATGGAAAGGCTGGAAAAGAAGAATTTTATAGgtactct atgcaaaaagattgtgtgttaacttttatg

# Figure 1E

# Exon 17

ttatttgttcagGGCTCTGTGTGACACTCCAGGTGTGGATCCAAAGCTTATTTCTAGAATTT GGGTTTATAATCACTATAGATGGATCATATGGAAACTGGCAGCTATGGAATGTGCC TTTCCTAAGGAATTTGCTAATAGATGCCTAAGCCCAGAAAGGGTGCTTCTTCAACTA AAATACAGgcaagtttaaagcatt ...

# Exon 18

ttttgttttcacttttagaTATGATACGGAAATTGATAGAAGCAGAAGATCGGCTATAAAAAAGA TAATGGAAAGGGATGACACAGCTGCAAAAACACTTGTTCTCTGTGTTTCTGACATA ATTTCATTGAGCGCAAATATATCTGAAACTTCTAGCAATAAAACTAGTAGTGCAGAT ACCCAAAAAGTGGCCATTATTGAACTTACAGATGGGTGGTATGCTGTTAAGGCCCA GTTAGATCCTCCCCTCTTAGCTGTCTTAAAGAATGGCAGACTGACAGTTGGTCAGA AGATTATTCTTCATGGAGCAGAACTGGTGGGCTCTCCTGATGCCTGTACACCTCTT GAAGCCCCAGAATCTCTTATGTTAAAGgtaaatt

## Exon 19

#### Exon 20

#### Exon 21

agtitagtgaattaataatcctttigttticttagAAAACACAACAAAACCATATTTACCATCACGTGCAC TAACAAGACAGCAAGTTCGTGCTTTGCAAGATGGTGCAGAGCTTTATGAAGCAGTG AAGAATGCAGCAGACCCAGCTTACCTTGAGgtgagagagtaagaggacatataatgag

#### Exon 22

tttttattccaatatcttaaatggtcacagGGTTATTTCAGTGAAGAGCAGTTAAGAGCCTTGAATAA TCACAGGCAAATGTTGAATGATAAGAAACAAGCTCAGATCCAGTTGGAAATTAGGA AGGCCATGGAATCTGCTGAACAAAAGGAACAAGGTTTATCAAGGGATGTCACAAC CGTGTGGAAGTTGCGTATTGTAAGCTATTCAAAAAAAAGAAAAAGATTCAGgtaagtatgt aaatgctttgttttta

## Exon 23

tctccaaacagTTATACTGAGTATTTGGCGTCCATCATCAGATTTATATTCTCTGTTAACA GAAGGAAAGAGATACAGAATTTATCATCTTGCAACTTCAAAATCTAAAAGTAAATCT GAAAGAGCTAACATACAGTTAGCAGCGACAAAAAAACTCAGTATCAACAACTACC Ggtacaaacctttcattgtaattttt

# Figure 1F

## Exon 24

## Exon 25

taacattcttttctttttttttccattctagGACTTGCCCCTTTCGTCTATTTGTCAGACGAATGTTACAA TTTACTGGCAATAAAGTTTTGGATAGACCTTAATGAGGACATTATTAAGCCTCATAT GTTAATTGCTGCAAGCAACCTCCAGTGGCGACCAGAATCCAAATCAGGCCTTCTTA CTTTATTTGCTGGAGATTTTTCTGTGTTTTCTGCTAGTCCAAAAGAGGGCCACTTTC AAGAGACATTCAACAAAATGAAAAATACTGTTGAGgtaaggtta

## Exon 26

ataaagcagcttttccacttattttcttagAATATTGACATACTTTGCAATGAAGCAGAAAACAAGCT TATGCATATACTGCATGCAAATGATCCCAAGTGGTCCACCCCAACTAAAGACTGTA CTTCAGGGCCGTACACTGCTCAAATCATTCCTGGTACAGGAAACAAGCTTCTGgtaa gttaatgtaaactcaaggaatattataag

#### Exon 27

# Figure 2A

#### Exon 2

taagtgcattttggtcttctgttttgcagACTTATTTACCAAGCATTGGAGGAATATCGTAGGTAAAA <u>ATG</u>CCTATTGGATCCAAAGAGGGCCAACATTTTTTGAAATTTTTAAGACACGCTGC AACAAAGCAGgtattgacaaattttatataac

#### Exon 3

gggattttttttttaaatagATTTAGGACCAATAAGTCTTAATTGGTTTGAAGAACTTTCTTCAGAAGCTCCACCCTATAATTCTGAACCTGCAGAAGAATCTGAACATAAAAACAACAATTACGAACCAAACCTATTTAAAACTCCACAAAGGAAACCATCTTATAATCAGCTGGCTTCAACTCCAATAATATTCAAAGAGCAAGGGCTGACTCTGCCGCTGTACCAATCTCCTGTAAAAGAATTAGATAAAATTCAAATTAGACTTAGGtaagtaatgcaatatggtagactgggg

#### Exon 4

tcactgaattattgtactgtttcagGAAGGAATGTTCCCAATAGTAGACATAAAAGTCTTCGCACA GTGAAAACTAAAATGGATCAAGCAGATGATGTTTCCTGTCCACTTCTAAATTCTTGT CTTAGTGAAAGgtatgatgaagctattatattaaaa

#### Exon 5

agggatttgctttgttttattttagTCCTGTTGTTCTACAATGTACACATGTAACACCACAAAGAGATAAAGTCAGgtatgattaaaaacaatgctttttattctt

#### Exon 6

ttaacaattttcccctttttttacccccagTGGTATGTGGGAGTTTGTTTCATACACCAAAGTTTGTGAAAGgtaaatatt

#### Exon 7

## Exon 8

ttttatcttacagTCAGAAATGAAGAAGCATCTGAAACTGTATTTCCTCATGATACTACTGC Tgtaagtaaatatgacattgattagact

#### Exon 9

## Exon 10

# Figure 2B

AGAAGCCCTTTGAGAGTGGAAGTGACAAAATCTCCAAGGAAGTTGTACCGTCTTTG GCCTGTGAATGGTCTCAACTAACCCTTTCAGGTCTAAATGGAGCCCAGATGGAGAA AATACCCCTATTGCATATTTCTTCATGTGACCAAAATATTTCAGAAAAAGACCTATTA GACACAGAGAACAAAGAAGAAGATTTTCTTACTTCAGAGAATTCTTTGCCACGT ATTTCTAGCCTACCAAAATCAGAGAAGCCATTAAATGAGGAAACAGTGGTAAATAA GAGAGATGAAGACACATCTTGAATCTCATACAGACTGCATTCTTGCAGTAAAGC AGGCAATATCTGGAACTTCTCCAGTGGCTTCTTCATTTCAGGGTATCAAAAAGTCTA TATTCAGAATAAGAGAATCACCTAAAGAGACTTTCAATGCAAGTTTTTCAGGTCATA TGACTGATCCAAACTTTAAAAAAGAAACTGAAGCCTCTGAAAGTGGACTGGAAATA CATACTGTTTGCTCACAGAAGGAGGACTCCTTATGTCCAAATTTAATTGATAATGGA AGCTGGCCAGCCACCACACAGAATTCTGTAGCTTTGAAGAATGCAGGTTTAAT ATCCACTTTGAAAAAGAAAACAAATAAGTTTATTTATGCTATACATGATGAAACATCT TATAAAGGAAAAAAATACCGAAAGACCAAAAATCAGAACTAATTAACTGTTCAGCC CAGTTTGAAGCAAATGCTTTTGAAGCACCACTTACATTTGCAAATGCTGATTCAGGt acctctgtct

## Exon 11

tttgtgtttttatgtttagGTTTATTGCATTCTTCTGTGAAAAGAAGCTGTTCACAGAATGATTCT GĂĂGAAČCAĂCTTTGTCCTTAACTAGCTCTTTTGGGACAATTCTGAGGAAATGTTCT AGAAATGAAACATGTTCTAATAATACAGTAATCTCTCAGGATCTTGATTATAAAGAA GCAAAATGTAATAAGGAAAAACTACAGTTATTTATTACCCCAGAAGCTGATTCTCTG TCATGCCTGCAGGAAGGACAGTGTGAAAATGATCCAAAAAGCAAAAAGTTTCAGA TATAAAAGAAGAGGTCTTGGCTGCAGCATGTCACCCAGTACAACATTCAAAAGTGG AATACAGTGATACTGACTTTCAATCCCAGAAAAGTCTTTTATATGATCATGAAAATG CCAGCACTCTTATTTTAACTCCTACTTCCAAGGATGTTCTGTCAAACCTAGTCATGA TTTCTAGAGGCAAAGAATCATACAAAATGTCAGACAAGCTCAAAGGTAACAATTATG CTTTAAATGAAAATTATAAAAACGTTGAGCTGTTGCCACCTGAAAAATACATGAGAG TAGCATCACCTTCAAGAAAGGTACAATTCAACCAAAACACAAATCTAAGAGTAATCC AAAAAAATCAAGAAGAAACTACTTCAATTTCAAAAATAACTGTCAATCCAGACTCTG AAGAACTTTCTCAGACAATGAGAATAATTTTGTCTTCCAAGTAGCTAATGAAAGGA ATAATCTTGCTTTAGGAAATACTAAGGAACTTCATGAAACAGACTTGACTTGTGTAA ACGAACCCATTTTCAAGAACTCTACCATGGTTTTATATGGAGACACAGGTGATAAAC AAGCAACCCAAGTGTCAATTAAAAAAGATTTGGTTTATGTTCTTGCAGAGGAGAAC AAAAATAGTGTAAAGCAGCATATAAAAATGACTCTAGGTCAAGATTTAAAATCGGAC ATCTCCTTGAATATAGATAAAATACCAGAAAAAAAATAATGATTACATGAACAAATGG GCAGGACTCTTAGGTCCAATTTCAAATCACAGTTTTGGAGGTAGCTTCAGAACAGC TTCAAATAAGGAAATCAAGCTCTCTGAACATAACATTAAGAAGAGCAAAATGTTCTT CAAAGATATTGAAGAACAATATCCTACTAGTTTAGCTTGTGTGAAATTGTAAATAC CTTGGCATTAGATAATCAAAAGAAACTGAGCAAGCCTCAGTCAATTAATACTGTATC TGCACATTTACAGAGTAGTGTAGTTGTTCTGATTGTAAAAATAGTCATATAACCCC TCAGATGTTATTTCCAAGCAGGATTTTAATTCAAACCATAATTTAACACCTAGCCAA AAGGCAGAAATTACAGAACTTTCTACTATATTAGAAGAATCAGGAAGTCAGTTTGAA TTTACTCAGTTTAGAAAACCAAGCTACATATTGCAGAAGAGTACATTTGAAGTGCCT GAAAACCAGATGACTATCTTAAAGACCACTTCTGAGGAATGCAGAGATGCTGATCT 

# Figure 2C

AAGGTACAGTTGAAATTAAACGGAAGTTTGCTGGCCTGTTGAAAAATGACTGTAAC AAAAGTGCTTCTGGTTATTTAACAGATGAAAATGAAGTGGGGTTTAGGGGCTTTTAT TCTGCTCATGGCACAAAACTGAATGTTTCTACTGAAGCTCTGCAAAAAGCTGTGAA ACTGTTTAGTGATATTGAGAATATTAGTGAGGAAACTTCTGCAGAGGTACATCCAAT AAGTTTATCTTCAAGTAAATGTCATGATTCTGTTGTTTCAATGTTTAAGATAGAAAAT CATAATGATAAAACTGTAAGTGAAAAAAAAAATAATAAATGCCAACTGATATTACAAAATA ATATTGAAATGACTACTGGCACTTTTGTTGAAGAAATTACTGAAAATTACAAGAGAA ATACTGAAAATGAAGATAACAAATATACTGCTGCCAGTAGAAATTCTCATAACTTAG AATTTGATGGCAGTGATTCAAGTAAAAATGATACTGTTTGTATTCATAAAGATGAAA CGGACTTGCTATTTACTGATCAGCACAACATATGTCTTAAATTATCTGGCCAGTTTA TGAAGGAGGAAACACTCAGATTAAAGAAGATTTGTCAGATTTAACTTTTTTGGAAG TTGCGAAAGCTCAAGAAGCATGTCATGGTAATACTTCAAATAAAGAACAGTTAACT GCTACTAAAACGGAGCAAAATATAAAAGATTTTGAGACTTCTGATACATTTTTTCAG ACTGCAAGTGGGAAAAATATTAGTGTCGCCAAAGAGTCATTTAATAAAATTGTAAAT TTCTTTGATCAGAAACCAGAAGAATTGCATAACTTTTCCTTAAATTCTGAATTACATT CTGACATAAGAAAGAACAAAATGGACATTCTAAGTTATGAGGAAACAGACATAGTT AAACACAAAATACTGAAAGAAAGTGTCCCAGTTGGTACTGGAAATCAACTAGTGAC CTTCCAGGGACAACCCGAACGTGATGAAAAGATCAAAGAACCTACTCTGTTGGGTT TTCATACAGCTAGCGGGAAAAAAGTTAAAATTGCAAAGGAATCTTTGGACAAAGTG AAAAACCTTTTTGATGAAAAAGAGCAAGGTACTAGTGAAATCACCAGTTTTAGCCAT CAATGGGCAAAGACCCTAAAGTACAGAGGGCCTGTAAAGACCTTGAATTAGCAT GTGAGACCATTGAGATCACAGCTGCCCCAAAGTGTAAAGAAATGCAGAATTCTCTC AATAATGATAAAAACCTTGTTTCTATTGAGACTGTGGTGCCACCTAAGCTCTTAAGT GATAATTTATGTAGACAAACTGAAAATCTCAAAAACATCAAAAAGTATCTTTTTGAAAG TTAAAGTACATGAAAATGTAGAAAAAGAAAAGCAAAAAAGTCCTGCAACTTGTTACA CAAATCAGTCCCTTATTCAGTCATTGAAAATTCAGCCTTAGCTTTTTACACAAGTT GTAGTAGAAAAACTTCTGTGAGTCAGACTTCATTACTTGAAGCAAAAAAATGGCTTA GAGAAGGAATATTTGATGGTCAACCAGAAAGAATAAATACTGCAGATTATGTAGGA AATTATTTGTATGAAAATAATTCAAACAGTACTATAGCTGAAAAATGACAAAAATCATC TCTCCGAAAACAAGATACTTATTTAAGTAACAGTAGCATGTCTAACAGCTATTCCT ACCATTCTGATGAGGTATATAATGATTCAGGATATCTCTCAAAAAATAAACTTGATT CTGGTATTGAGCCAGTATTGAAGAATGTTGAAGATCAAAAAAACACTAGTTTTTCCA AAGTAATATCCAATGTAAAAGATGCAAATGCATACCCACAAACTGTAAATGAAGATA TTTGCGTTGAGGAACTTGTGACTAGCTCTTCACCCTGCAAAAATAAAAATGCAGCC ATTAAATTGTCCATATCTAATAGTAATAATTTTGAGGTAGGGCCACCTGCATTTAGG ATAGCCAGTGGTAAAATCGTTTGTGTTTCACATGAAACAATTAAAAAAAGTGAAAGAC ATATTTACAGACAGTTTCAGTAAAGTAATTAAGGAAAACAACGAGAATAAATCAAAA ATTTGCCAAACGAAAATTATGGCAGGTTGTTACGAGGCATTGGATGATTCAGAGGA TATTCTTCATAACTCTCTAGATAATGATGAATGTAGCACGCATTCACATAAGGTTTTT GCTGACATTCAGAGTGAAGAAATTTTACAACATAACCAAAATATGTCTGGATTGGA GAAAGTTTCTAAAATATCACCTTGTGATGTTAGTTTGGAAACTTCAGATATATGTAAA TGTAGTATAGGGAAGCTTCATAAGTCAGTCTCATCTGCAAATACTTGTGGGATTTTT AGCACAGCAAGTGGAAAATCTGTCCAGGTATCAGATGCTTCATTACAAAACGCAAG ACAAGTGTTTCTGAAATAGAAGATAGTACCAAGCAAGTCTTTTCCAAAGTATTGTT CTCCAGAACATTTAATATCCCAAAAAGGCTTTTCATATAATGTGGTAAATTCATCTG

# Figure 2D

CTTTCTCTGGATTTAGTACAGCAAGTGGAAAGCAAGTTTCCATTTTAGAAAGTTCCT TACACAAAGTTAAGGGAGTGTTAGAGGAATTTGATTTAATCAGAACTGAGCATAGT CTTCACTATTCACCTACGTCTAGACAAAATGTATCAAAAATACTTCCTCGTGTTGAT AAGAGAAACCCAGAGCACTGTGTAAACTCAGAAATGGAAAAAACCTGCAGTAAAGA ATTTAAATTATCAAATAACTTAAATGTTGAAGGTGGTTCTTCAGAAAATAATCACTCT ATTAAAGTTTCTCCATATCTCTCAATTTCAACAAGACAACAACAGTTGGTATTAG GAACCAAAGTCTCACTTGTTGAGAACATTCATGTTTTGGGAAAAGAACAGGCTTCA CCTAAAAACGTAAAAATGGAAATTGGTAAAACTGAAACTTTTTCTGATGTTCCTGTG AAAACAAATATAGAAGTTTGTTCTACTTACTCCAAAGATTCAGAAAACTACTTTGAAA CAGAAGCAGTAGAAATTGCTAAAGCTTTTATGGAAGATGATGAACTGACAGATTCT AAACTGCCAAGTCATGCCACACATTCTCTTTTTACATGTCCCGAAAATGAGGAAATG aagtgttcatttttacctttcgtgttgccaatca

# Exon 12

aaaacatatatgaaatatttctttttagGAGAACCCTCAATCAAAAGAAACTTATTAAATGAATTTG ACAGGATAATAGAAAATCAAGAAAAATCCTTAAAGGCTTCAAAAAGCACTCCAGAT Ggtaaaattagctttttatttata

# Exon 13

aatatgtaatataaaataattgtttcctagGCACAATAAAAGATCGAAGATTGTTTATGCATCATGT TTCTTTAGAGCCGATTACCTGTGTACCCTTTCGgtaagacatgtttaaatttttctaa

#### Exon 14

ccccattgcagCACAACTAAGGAACGTCAAGAGATACAGAATCCAAATTTTACCGCACC TGGTČAAĞAATTTCTGTCTAAATCTCATTTGTATGAACATCTGACTTTGGAAAAATCT TCAAGCAATTTAGCAGTTTCAGGACATCCATTTTATCAAGTTTCTGCTACAAGAAAT GAAAAATGAGACACTTGATTACTACAGGCAGACCAAACCAAAGTCTTTGTTCCACC TTTTAAAACTAAATCACATTTTCACAGAGTTGAACAGTGTGTTAGGAATATTAACTTG GAGGAAAACAGACAAAGCAAAACATTGATGGACATGGCTCTGATGATAAAAAA TAAGATTAATGACAATGAGATTCATCAGTTTAACAAAAACAACTCCAATCAAGCAGC AGCTGTAACTTTCACAAAGTGTGAAGAAGAACCTTTAGgtattgtatgacaatttgtgtgatgaatt

#### Exon 15

tttttgctaagtatttattctttgatagATTTAATTACAAGTCTTCAGAATGCCAGAGATATACAGGAT ATĞCGĂATTAAGĂAGĂAACAAAGGCAACGCGTCTTTCCACAGCCAGGCAGTCTGTA TCTTGCAAAAACATCCACTCTGCCTCGAATCTCTCTGAAAGCAGCAGTAGGAGGCC AAGTTCCCTCTGCGTGTTCTCATAAACAGgtatgtgt

## Exon 16

 $tittictttttttgtgtgtgtttattttgtgtag {\tt CTGTATACGTATGGCGTTTCTAAACATTGCATAAAAATTA}\\$ ACAGCĂĂĂĂTGCĂĞAĞTCTTTTCAGTTTCACACTGAAGATTATTTTGGTAAGGAAA GTTTATGGACTGGAAAAGGAATACAGTTGGCTGATGGTGGATGGCTCATACCCTCC AATGATGGAAAGGCTGGAAAAGAAGAATTTTATAGgtactctatgcaaaaagattgtgtgttaactttt atg

# Figure 2E

Exon 17

ttatttgttcagGGCTCTGTGTGACACTCCAGGTGTGGATCCAAAGCTTATTTCTAGAATTT GGĞTTTĂTAATCACTATAGATGGATCATATGGAAACTGGCAGCTATGGAATGTGCC TTTCCTAAGGAATTTGCTAATAGATGCCTAAGCCCAGAAAGGGTGCTTCTTCAACTA AAATACAGgcaagtttaaagcatt

Exon 18

TAATGGAAAGGGATGACACAGCTGCAAAAACACTTGTTCTCTGTGTTTCTGACATA ATTTCATTGAGCGCAAATATATCTGAAACTTCTAGCAATAAAACTAGTAGTGCAGAT ACCCAAAAGTGGCCATTATTGAACTTACAGATGGGTGGTATGCTGTTAAGGCCCA GTTAGATCCTCCCCTCTTAGCTGTCTTAAAGAATGGCAGACTGACAGTTGGTCAGA AGATTATTCTTCATGGAGCAGAACTGGTGGGCTCTCCTGATGCCTGTACACCTCTT GAAGCCCCAGAATCTCTTATGTTAAAGgtaaatt

Exon 19

taaatcaatatttattaatttgtccagATTTCTGCTAACAGTACTCGGCCTGCTCGCTGGTATAC CAAACTTGGATTCTTCCTGACCCTAGACCTTTTCCTCTGCCCTTATCATCGCTTTT CAGTGATGGAGGAAATGTTGGTTGTTGATGTAATTATTCAAAGAGCATACCCTAT ACAGgtatgatgtattcttgaaactta

Exon 20

tttggtgtgtgtaacacattattacagTGGATGGAGAAGACATCATCTGGATTATACATATTTCGCAĂŤĞĂĂĞAĞAĞĞAAĞĂAAAĞĞAAĞCAĞCAAAATATĞTĞĞAĞĞCCCAACAAAĞĀ GACTAGAAGCCTTATTCACTAAAATTCAGGAGGAATTTGAAGAACATGAAGgtaaaatt agttatatggtacacattgttatttc

Exon 21

agtttagtgaattaataatccttttgttttcttagAAAACACAACAAAACCATATTTACCATCACGTGCAC TAACAAGACAGCAAGTTCGTGCTTTGCAAGATGGTGCAGAGCTTTATGAAGCAGTG AAGAATGCAGCAGACCCAGCTTACCTTGAGgtgagaggtaagaggacatataatgag

Exon 22

tttttattccaatatcttaaatggtcacagGGTTATTTCAGTGAAGAGCAGTTAAGAGCCTTGAATAA TCACAGGCAAATGTTGAATGATAAGAAACAAGCTCAGATCCAGTTGGAAATTAGGA AGgCCATGGAATCTGCTGAACAAAAGGAACAAGGTTTATCAAGGGATGTCACAACC GTGTGGAAGTTGCGTATTGTAAGCTATTCAAAAAAAAGAAAAAGATTCAGgtaagtatgta aatgctttgttttta

Exon 23

tctccaaacagTTATACTGAGTATTTGGCGTCCATCATCAGATTTATATTCTCTGTTAACA GAAGGAAAGGAATTTATCATCTTGCAACTTCAAAATCTAAAAGTAAATCT GAAAGAGCTAACATACAGTTAGCAGCGACAAAAAAAACTCAGTATCAACAACTACC Gotacaaacctttcattgtaattttt

# Figure 2F

Exon 24

gaatttttgttttgttttctgtagGTTTCAGATGAAATTTTATTTCAGATTTACCAGCCACGGGAGC CCCTTCACTTCAGCAAATTTTTAGATCCAGACTTTCAGCCATCTTGTTCTGAGGTGG ACCTAATAGGATTTGTCGTTTCTGTTGTGAAAAAAAACAGgtaatgcacaatatagttaatttttttat tgattcttttaaaaaaacattgtct

Exon 25

taacattcttttcttttttttccattctagGACTTGCCCCTTTCGTCTATTTGTCAGACGAATGTTACAA TTTACTGGCAATAAAGTTTTGGATAGACCTTAATGAGGACATTATTAAGCCTCATAT GTTAATTGCTGCAAGCAACCTCCAGTGGCGACCAGAATCCAAATCAGGCCTTCTTA CTTTATTTGCTGGAGATTTTTCTGTGTTTTCTGCTAGTCCAAAAGAGGGCCACTTTC AAGAGACATTCAACAAAATGAAAAATACTGTTGAGgtaaggtta

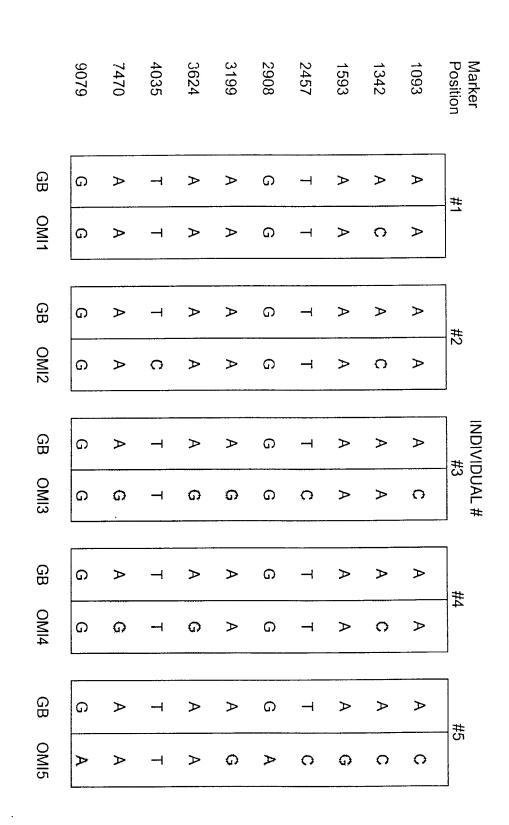
Exon 26

ataaagcagcttttccacttattttcttagAATATTGACATACTTTGCAATGAAGCAGAAAACAAGCT TATGCATATACTGCATGCĂAATGATCCCAAGTGGTCCACCCCAACTAAAGACTGTA CTTCAGGGCCGTACACTGCTCAAATCATTCCTGGTACAGGAAACAAGCTTCTGgtaa gttaatgtaaactcaaggaatattataag

Exon 27

tacgttttcatttttttatcagATGTCTTCTCCTAATTGTGAGATATATTATCAAAGTCCTTTATCA CTTTGTATGGCCAAAAGGAAGTCTGTTTCCACACCTGTCTCAGCCCAGATGACTTC AAAGTCTTGTAAAGGGGAGAAAGAGATTGATGACCAAAAGAACTGCAAAAAGAGAA GAGCCTTGGATTTCTTGAGTAGACTGCCTTTACCTCCACCTGTTAGTCCCATTTGTA CATTTGTTTCTCCGGCTGCACAGAAGGCATTTCAGCCACCAAGGAGTTGTGGCAC CAAATACGAAACACCCATAAAGAAAAAAGAACTGAATTCTCCTCAGATGACTCCATT TAAAAAATTCAATGAAATTTCTCTTTTGGAAAGTAATTCAATAGCTGACGAAGAACTT GCATTGATAAATACCCAAGCTCTTTTGTCTGGTTCAACAGGAGAAAAAACAATTTATA TCTGTCAGTGAATCCACTAGGACTGCTCCCACCAGTTCAGAAGATTATCTCAGACT GAAACGACGTTGTACTACATCTCTGATCAAAGAACAGGAGAGTTCCCAGGCCAGTA CGGAAGAATGTGAGAAAAATAAGCAGGACACAATTACAACTAAAAAATATATCTAA GCATTTGCAAAGGCGACAATAAATTATTGACGCTTAACCTTTCCAGTTTATAAGACT **GGA** 

# FIGURE 3



VERIFIED STATEMENT CLAIMING SMALL ENTITY STATUS	Docket No:	
(37 CFR 1.9(f) & 1.27(c)) - SMALL BUSINESS CONCERN	05371.31.US02	
	PA-	

Applicant or Patentee:

Patricia D. Murphy; Marga B. White; Mark B. Rabin; Sheri J. Olson; Matthew Yoshikawa; Geoffrey

M. Jackson; Tara Eskandari; Brenda Schryer; and Michael Park.

Serial or Patent No.:

To be assigned.

Filed or Issued:

Herewith

Title:

NOVEL CODING SEQUENCE HAPLOTYPES OF THE HUMAN BRCA2 GENE

I hereby declare that I am

the owner of the small business concern identified below:

XXX an official of the small business concern empowered to act on behalf of the concern identified below:

NAME OF SMALL BUSINESS CONCERN:

Oncormed, Inc.

ADDRESS OF SMALL BUSINESS CONCERN: 205 Perry Parkway Gaithersburg, MD 20877

I hereby declare that the above-identified small business concern qualifies as a small business concern as defined in 13 CFR 121.12, and reproduced in 37 CFR 1.9(d), for purposes of paying reduced fees to the United States Patent and Trademark Office, in that the number of employees of the concern, including those of its affiliates, does not exceed 500 persons. For purposes of this statement, (1) the number of employees of the business concern is the average over the pervious fiscal year of the concern of the persons employed on a full-time, part-time or temporary basis during each of the pay periods of the fiscal year, and (2) concerns are affiliates of each other when either, directly or indirectly, one concern controls of has the power to control the other, or a third party or parties controls or has the power to control both.

I hereby declare that rights under contract or law have been conveyed to and remain with the small business concern identified above with regard to the invention described in:

XXXX_	the specification filed herewith with title as listed above.
	the application identified above.
	the patent identified above.

If the rights held by the above identified small business concern are not exclusive, each individual, concern or organization having rights in the invention must file separate verified statements averring to their status as small entities, and no rights to the invention are held by any person, other than the inventor, who would not qualify as an independent inventor under 37 CFR 1.9(c) if that person made the invention, or by any concern which would not qualify as a small business concern under 37 CFR 1.9(d), or a nonprofit organization under 37 CFR 1.9(e).

	Each person, concern or organization having any rights in the invention is listed below.
XXX	No such person, concern or organization exists.
	Each such person, concern or organization is listed below.

Separate verified statements are required from each named person, concern or organization having rights to the invention averring to their status as small entities. (37 CFR 1.27)

I acknowledge the duty to file, in this application or patent, notification of any change in status resulting in loss of entitlement to small entity status prior to paying, or at the time of paying, the earliest of the issue fee or any maintenance fees due after the date on which status as a small entity is no longer appropriate. (37 CFR 1.28(b)).

I hereby declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under section 1001 of Title 18 of the United States Code, and that such willful false statements may jeopardize the validity of the application, and patent issuing thereon, or any patent to which this verified statement is directed.

NAME OF PERSON SIGNING:

President & C.O.O.

TITLE OF PERSON IF OTHER THAN OWNER: ADDRESS OF PERSON SIGNING:

205 Perry Parkway Gaithersburg, MD 20877

DOUG DOLGINOW, M.D.

DATE 5/13/94

# Combined Declaration and Power of Attorney for Patent Application

Docket Number: 5371.31.US02

As a below named inventor, I hereby declare that:

My residence, post office address and citizenship are as stated below next to my name.

I believe I am an original, first and joint inventor (if plural names are listed below) of the subject matter that is claimed and for which a patent is sought on the invention entitled **NOVEL CODING SEQUENCE HAPLOTYPES OF THE HUMAN BRCA2 GENE**, the specification of which is attached hereto unless the following box is checked:

was filed on	Herewith		;				
as United States	Application N	Number or PCT	International	Application Number	T	o Be Assigned	_; and
was amended on			_(if applicable)	).			

I hereby state that I have reviewed and understand the contents of the above identified specification, including the claims, as amended by any amendment referred to above.

I acknowledge the duty to disclose information that is material to patentability as defined in 37 C.F.R. § 1.56.

I hereby claim foreign priority benefits under 35 U.S.C. § 119(a)-(d) or § 365(b) of any foreign application(s) for patent or inventor's certificate, or § 365(a) of any PCT international application, which designated at least one country other than the United States listed below, and have also identified below any foreign application for patent or inventor's certificate, or PCT international application having a filing date before that of the application on which priority is claimed.

OR FOREIGN APPLICATE Application No.	Country	(Day/Month/Year/Filed)	Priority Claimed
			Yes No

I hereby claim the benefit under 35 U.S.C. §119(e) of any United States provisional application(s) listed below.

Application No.	Filing Date
60/055,784	August 15, 1997
60/064,926	November 7, 1997
60/065,367	November 12, 1997

I hereby claim the benefit under 35 U.S.C. § 120 of any United States application(s), or under § 365(c) of any PCT international application designating the United States, listed below and, insofar as the subject matter of each of the claims of this application is not disclosed in the prior United States or PCT international application in the manner provided by the first paragraph of 35 U.S.C. § 112, I acknowledge the duty to disclose information that is material to patentability as defined in 37 C.F.R. § 1.56 that became available between the filing date of the prior application and the national or PCT international filing date of this application.

Application No.	Filing Date	(Status – patented, pending, abandoned)

I hereby appoint the following attorney(s) and/or agent(s) to prosecute this application and to transact all business in the Patent and Trademark Office connected therewith:

Jeffrey I. Auerbach, Reg. No. 32,680 Melvin L. Barnes, Jr. Reg. No. 38,375 Thane Bauz, Reg. No. P41,604 Michael J. Bell, Reg. No. 39,604 John A. Bendrick, Reg. No. P41,612 Mark R. Buscher, Reg. No. 35,006 Celine T. Callahan, Reg. No. 34,301 Cono A. Carrano, Reg. No. 39,623 James F. Davis, Reg. No. 21,072 Thomas M. Dunham, Reg. No. 39,965 Joel M. Freed, Reg. No. 25,101 Vernon Randall Gard, Reg. No. 33,886

Alan M. Grimaldi, Reg. No. 26,599 Alexander J. Hadjis, Reg. No. 36,540 Albert P. Halluin, Reg. No. 25,227 Michael N. Haynes, Reg. No. 40,014 Rouget F. Henschel, Reg. No. 39,221 Leslie L. Jacobs, Jr., Reg. No. 40,659 Richard H. Kjeldgaard, Reg. No. 30,186 Joseph P. Lavelle, Reg. No. 31,036 David R. Marsh, Reg. No. 41,408 Kevin W. McCabe, Reg. No. 41,182 Joseph A. Micallef, Reg. No. 39,772 Anthony D. Miller, Reg. No. 34,394 Karen L. Nicastro, Reg. No. 35,968
Bradley J. Olson, Reg. No. 40,750
Russell O. Paige, Reg. No. P40,758
Stephen J. Pentlicki, Reg. No. 40,125
Andrew Y. Piatnicia, Reg. No. 40,772
Andrea G. Reister, Reg. No. 36,253
Stephen J. Rosenman, Reg. No. 29,209
David P. Ruschke, Reg. No. 40,151
Timothy L. Scott, Reg. No. 37,931
Anthony W. Shaw, Reg. No. 30,104
J. David Smith, Reg. No. 39,839
Michael J. Songer, Reg. No. 39,841
R. Thomas Gallages
Reg. No. 32, 692
John E. Tarcza
Reg. No. 33,683

Send Correspondence to:

## Albert P. Halluin HOWREY & SIMON

Box No. 34 1299 Pennsylvania Avenue, N.W. Washington, D.C. 20004-2402 Facsimile: (202) 383-7195

Direct Telephone Calls to: (202) 783-0800

I hereby declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code and that such willful false statements may jeopardize the validity of the application or any patent issued thereon.

FULL NAME OF SOLE OR FIRST INVENTOR:	CITIZENSHIP:
Patricia D. Murphy	USA
RESIDENCE: Slingerlands, New York, USA	DATE: 578/98
Singerlands, New York, OSA	
POST OFFICE ADDRESS:	INVENTOR'S SIGNATURE:
205 Perry Parkway, Gaithersburg, Maryland 20877, USA	Janua & murphy
FULL NAME OF SECOND INVENTOR:	CITIZENSHIP:
Marga B. White	USA
RESIDENCE:	DATE:
	5,13/95
Frederick, Maryland, USA	)   13
POST OFFICE ADDRESS:	INVENTOR'S SIGNATURE:
205 Perry Parkway, Gaithersburg, Maryland 20877, USA	Mary Bille Come
,	
FULL NAME OF THIRD INVENTOR:	CITIZENSHIP:
Mark B. Rabin	USA
RESIDENCE:	DATE: /
Rockville, Maryland, USA	5/13/98
POST OFFICE ADDRESS:	INVENTOR'S SIGNATURE:
205 Perry Parkway, Gaithersburg, Maryland 20877, USA	11 Care Pull
	<del>,</del>
<u> </u>	

(Supply similar information and signature for subsequent joint inventors, if any)

O
=
=
١,٠٠٠
=
U
Ų
. <b>"</b>
Ü

FULL NAME OF FOURTH INVENTOR:	CITIZENSHIP USA
Sheri J. Olson	
RESIDENCE:	DATE: 5/13/98
Falls Church, Virginia, USA	4
POST OFFICE ADDRESS:	INVENTORISTICNATURE:
205 Perry Parkway, Gaithersburg, Maryland 20877, USA	The Man Care
FULL NAME OF FIFTH INVENTOR:	CITIZENSHIP:
Matthew Yoshikawa	USA
RESIDENCE ADDRESS:	DATE:
Germantown, Maryland, USA	
POST OFFICE ADDRESS:	INVENTOR'S SIGNATURE:
205 Perry Parkway, Gaithersburg, Maryland 20877, USA	
FULL NAME OF SIXTH INVENTOR:	CITIZENSHIP:
Geoffrey M. Jackson	USA
RESIDENCE ADDRESS:	DATE:
Beltsville, Maryland, USA	
POST OFFICE ADDRESS:	INVENTOR'S SIGNATURE:
205 Perry Parkway, Gaithersburg, Maryland 20877, USA	
and A strip A minimary Committee and James and	
FULL NAME OF SEVENTH INVENTOR:	CITIZENSHIP:
Tara Eskandari	USA
RESIDENCE ADDRESS:	
RESIDENCE ADDRESS.	DATE: 5-13-98
Rockville, Maryland, USA	
Rockville, Maryland, USA	
POST OFFICE ADDRESS:	
	INVENTOR'S SIGNATURE:
POST OFFICE ADDRESS: 205 Perry Parkway, Gaithersburg, Maryland 20877, USA	INVENTOR'S SIGNATURE:
POST OFFICE ADDRESS: 205 Perry Parkway, Gaithersburg, Maryland 20877, USA  FULL NAME OF EIGHTH INVENTOR:	INVENTOR'S SIGNATURE:  CITIZENSHIP:
POST OFFICE ADDRESS: 205 Perry Parkway, Gaithersburg. Maryland 20877, USA  FULL NAME OF EIGHTH INVENTOR: Brenda Schryer	CITIZENSHIP: USA
POST OFFICE ADDRESS: 205 Perry Parkway, Gaithersburg, Maryland 20877, USA  FULL NAME OF EIGHTH INVENTOR: Brenda Schryer RESIDENCE ADDRESS:	CITIZENSHIP: USA DATE:
POST OFFICE ADDRESS: 205 Perry Parkway, Gaithersburg, Maryland 20877, USA  FULL NAME OF EIGHTH INVENTOR: Brenda Schryer  RESIDENCE ADDRESS: Bel Air, Maryland, USA	CITIZENSHIP: USA DATE: 5-16-98
POST OFFICE ADDRESS: 205 Perry Parkway, Gaithersburg, Maryland 20877, USA  FULL NAME OF EIGHTH INVENTOR: Brenda Schryer  RESIDENCE ADDRESS: Bel Air, Maryland, USA  POST OFFICE ADDRESS:	CITIZENSHIP: USA DATE: 5-16-98
POST OFFICE ADDRESS: 205 Perry Parkway, Gaithersburg, Maryland 20877, USA  FULL NAME OF EIGHTH INVENTOR: Brenda Schryer  RESIDENCE ADDRESS: Bel Air, Maryland, USA	CITIZENSHIP: USA DATE:
POST OFFICE ADDRESS: 205 Perry Parkway, Gaithersburg, Maryland 20877, USA  FULL NAME OF EIGHTH INVENTOR: Brenda Schryer  RESIDENCE ADDRESS: Bel Air, Maryland, USA  POST OFFICE ADDRESS: 205 Perry Parkway, Gaithersburg, Maryland 20877, USA	INVENTOR'S SIGNATURE:  CITIZENSHIP:  USA  DATE:  5-16-98  INVENTOR'S SIGNATURE:  Broad Carry Schare
POST OFFICE ADDRESS: 205 Perry Parkway, Gaithersburg, Maryland 20877, USA  FULL NAME OF EIGHTH INVENTOR: Brenda Schryer  RESIDENCE ADDRESS: Bel Air, Maryland, USA  POST OFFICE ADDRESS: 205 Perry Parkway, Gaithersburg, Maryland 20877, USA  FULL NAME OF NINTH INVENTOR:	CITIZENSHIP:  USA  DATE:  SOCIOLARIA  INVENTOR'S SIGNATURE:  BOOKELE SCHOOL  CITIZENSHIP:
POST OFFICE ADDRESS: 205 Perry Parkway, Gaithersburg, Maryland 20877, USA  FULL NAME OF EIGHTH INVENTOR: Brenda Schryer  RESIDENCE ADDRESS: Bel Air, Maryland, USA  POST OFFICE ADDRESS: 205 Perry Parkway, Gaithersburg, Maryland 20877, USA  FULL NAME OF NINTH INVENTOR: Michael Park	INVENTOR'S SIGNATURE:  CITIZENSHIP: USA  DATE:  5-16-98  INVENTOR'S SIGNATURE:  Boods Schaue  CITIZENSHIP: USA
POST OFFICE ADDRESS: 205 Perry Parkway, Gaithersburg, Maryland 20877, USA  FULL NAME OF EIGHTH INVENTOR: Brenda Schryer  RESIDENCE ADDRESS: Bel Air, Maryland, USA  POST OFFICE ADDRESS: 205 Perry Parkway, Gaithersburg, Maryland 20877, USA  FULL NAME OF NINTH INVENTOR: Michael Park  RESIDENCE ADDRESS:	INVENTOR'S SIGNATURE:  CITIZENSHIP:  USA  DATE:  5-16-98  INVENTOR'S SIGNATURE:  BOOKELE STATES  CITIZENSHIP:
POST OFFICE ADDRESS: 205 Perry Parkway, Gaithersburg, Maryland 20877, USA  FULL NAME OF EIGHTH INVENTOR: Brenda Schryer  RESIDENCE ADDRESS: Bel Air, Maryland, USA  POST OFFICE ADDRESS: 205 Perry Parkway, Gaithersburg, Maryland 20877, USA  FULL NAME OF NINTH INVENTOR: Michael Park	INVENTOR'S SIGNATURE:  CITIZENSHIP: USA  DATE:  5-16-98  INVENTOR'S SIGNATURE:  Boods Schaue  CITIZENSHIP: USA
POST OFFICE ADDRESS: 205 Perry Parkway, Gaithersburg, Maryland 20877, USA  FULL NAME OF EIGHTH INVENTOR: Brenda Schryer  RESIDENCE ADDRESS: Bel Air, Maryland, USA  POST OFFICE ADDRESS: 205 Perry Parkway, Gaithersburg, Maryland 20877, USA  FULL NAME OF NINTH INVENTOR: Michael Park  RESIDENCE ADDRESS:	INVENTOR'S SIGNATURE:  CITIZENSHIP:  USA  DATE:  5-16-98  INVENTOR'S SIGNATURE:  BOOKELE Schafe  CITIZENSHIP:  USA
POST OFFICE ADDRESS: 205 Perry Parkway, Gaithersburg. Maryland 20877, USA  FULL NAME OF EIGHTH INVENTOR: Brenda Schryer  RESIDENCE ADDRESS: Bel Air, Maryland, USA  POST OFFICE ADDRESS: 205 Perry Parkway, Gaithersburg, Maryland 20877, USA  FULL NAME OF NINTH INVENTOR: Michael Park  RESIDENCE ADDRESS: Rockville, Maryland, USA	CITIZENSHIP: USA  DATE:  SOCIOLAR SCHOOL  CITIZENSHIP: USA  DATE:  CITIZENSHIP: USA  DATE:  D

## Combined Declaration and Power of Attorney for Patent Application

Docket Number: 5371.31.US02

As a below named inventor, I hereby declare that:

My residence, post office address and citizenship are as stated below next to my name.

I believe I am an original, first and joint inventor (if plural names are listed below) of the subject matter that is claimed and for which a patent is sought on the invention entitled **NOVEL CODING SEQUENCE HAPLOTYPES OF THE HUMAN BRCA2 GENE**, the specification of which is attached hereto unless the following box is checked:

was filed on	Herewith	;		
as United States	Application	Number or PCT International Application Number	To Be Assigned; an	nd
was amended or	1	(if applicable).		

I hereby state that I have reviewed and understand the contents of the above identified specification, including the claims, as amended by any amendment referred to above.

I acknowledge the duty to disclose information that is material to patentability as defined in 37 C.F.R. § 1.56.

I hereby claim foreign priority benefits under 35 U.S.C. § 119(a)-(d) or § 365(b) of any foreign application(s) for patent or inventor's certificate, or § 365(a) of any PCT international application, which designated at least one country other than the United States listed below, and have also identified below any foreign application for patent or inventor's certificate, or PCT international application having a filing date before that of the application on which priority is claimed.

Application No.	Country	(Day/Month/Year/Filed)	Priority Claimed
			Yes No

I hereby claim the benefit under 35 U.S.C. §119(e) of any United States provisional application(s) listed below.

Application No.	Filing Date
60/055,784	August 15, 1997
60/064,926	November 7, 1997
60/065,367	November 12, 1997

I hereby claim the benefit under 35 U.S.C. § 120 of any United States application(s), or under § 365(c) of any PCT international application designating the United States, listed below and, insofar as the subject matter of each of the claims of this application is not disclosed in the prior United States or PCT international application in the manner provided by the first paragraph of 35 U.S.C. § 112, I acknowledge the duty to disclose information that is material to patentability as defined in 37 C.F.R. § 1.56 that became available between the filing date of the prior application and the national or PCT international filing date of this application.

Application No.	Filing Date	(Status – patented, pending, abandoned)

I hereby appoint the following attorney(s) and/or agent(s) to prosecute this application and to transact all business in the Patent and Trademark Office connected therewith:

Jeffrey I. Auerbach, Reg. No. 32,680 Melvin L. Barnes, Jr. Reg. No. 38,375 Thane Bauz, Reg. No. P41,604 Michael J. Bell, Reg. No. 39,604 John A. Bendrick, Reg. No. P41,612 Mark R. Buscher, Reg. No. 35,006 Celine T. Callahan, Reg. No. 34,301 Cono A. Carrano, Reg. No. 39,623 James F. Davis, Reg. No. 21,072 Thomas M. Dunham, Reg. No. 39,965 Joel M. Freed, Reg. No. 25,101 Vernon Randall Gard, Reg. No. 33,886

Send Correspondence to:

Alan M. Grimaldi, Reg. No. 26,599
Alexander J. Hadjis, Reg. No. 36,540
Albert P. Halluin, Reg. No. 25,227
Michael N. Haynes, Reg. No. 40,014
Rouget F. Henschel, Reg. No. 39,221
Leslie L. Jacobs, Jr., Reg. No. 40,659
Richard H. Kjeldgaard, Reg. No. 30,186
Joseph P. Lavelle, Reg. No. 31,036
David R. Marsh, Reg. No. 41,408
Kevin W. McCabe, Reg. No. 41,182
Joseph A. Micallef, Reg. No. 39,772
Anthony D. Miller, Reg. No. 34,394

## Albert P. Halluin HOWREY & SIMON

Box No. 34 1299 Pennsylvania Avenue, N.W. Washington, D.C. 20004-2402 Facsimile: (202) 383-7195 Karen L. Nicastro, Reg. No. 35,968
Bradley J. Olson, Reg. No. 40,750
Russell O. Paige, Reg. No. P40,758
Stephen J. Pentlicki, Reg. No. 40,125
Andrew Y. Piatnicia, Reg. No. 40,772
Andrea G. Reister, Reg. No. 36,253
Stephen J. Rosenman, Reg. No. 29,209
David P. Ruschke, Reg. No. 40,151
Timothy L. Scott, Reg. No. 37,931
Anthony W. Shaw, Reg. No. 30,104
J. David Smith, Reg. No. 39,839
Michael J. Songer, Reg. No. 39,841
R. Thomas Galleges
Reg. No. 32, 692
John E. Tarcza
Reg. No. 33, 683

Direct Telephone Calls to: (202) 783-0800

I hereby declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code and that such willful false statements may jeopardize the validity of the application or any patent issued thereon.

FULL NAME OF SOLE OR FIRST INVENTOR:	CITIZENSHIP:
Patricia D. Murphy	USA
RESIDENCE:	DATE:
Slingerlands, New York, USA	578/98
POST OFFICE ADDRESS:	INVENTOR'S SIGNATURE:
205 Perry Parkway, Gaithersburg, Maryland 20877, USA	Satrua & murphy
FULL NAME OF SECOND INVENTOR:	CITIZENSHIP:
Marga B. White	USA
RESIDENCE:	DATE:
Frederick, Maryland, USA	5/13/98
POST OFFICE ADDRESS:	INVENTOR'S SIGNATURE:
205 Perry Parkway, Gaithersburg, Maryland 20877, USA	mange but white
FULL NAME OF THIRD INVENTOR:	CITIZENSHIP:
Mark B. Rabin	USA
RESIDENCE:	DATE:
Rockville, Maryland, USA	5/13/98
POST OFFICE ADDRESS:	INVENTOR'S SIGNATURE:
205 Perry Parkway, Gaithersburg, Maryland 20877, USA	Illak Kile

(Supply similar information and signature for subsequent joint inventors, if any)

FULL NAME OF FOURTH INVENTOR:	CITIZENSHIP
Sheri J. Olson	USA
RESIDENCE:	DATE: 5/2/00
Falls Church, Virginia, USA	5/13/98
POST OFFICE ADDRESS:	INVENTOR'S SIGNATURE:
205 Perry Parkway, Gaithersburg, Maryland 20877, USA	Mon don Lison.
FULL NAME OF FIFTH INVENTOR:	CITIZENSHIP:
Matthew Yoshikawa	USA
RESIDENCE ADDRESS:	DATE: 5/18/98
Germantown, Maryland, USA	3/18/16
POST OFFICE ADDRESS:	INVENTOR'S SIGNATURE:
205 Perry Parkway, Gaithersburg, Maryland 20877, USA	Redden Joseph Ce
FULL NAME OF SIXTH INVENTOR:	CITIZENSHIP:
Geoffrey M. Jackson	USA
RESIDENCE ADDRESS:	DATE: 4/5/98
Beltsville, Maryland, USA	3/13/10
POST OFFICE ADDRESS:	INVENTOR'S SIGNATURE:
205 Perry Parkway, Gaithersburg, Maryland 20877, USA	the forhson
FULL NAME OF SEVENTH INVENTOR:	CITIZENSHIP:
Tara Eskandari	USA
RESIDENCE ADDRESS:	DATE: 5-13-98
Rockville, Maryland, USA	
POST OFFICE ADDRESS:	INVENTOR'S SIGNATURE:
205 Perry Parkway, Gaithersburg, Maryland 20877, USA	Tellarena
FULL NAME OF EIGHTH INVENTOR:	CITIZENSHIP:
Brenda Schryer	USA
RESIDENCE ADDRESS:	DATE:
Bel Air, Maryland, USA	
POST OFFICE ADDRESS:	INVENTOR'S SIGNATURE:
205 Perry Parkway, Gaithersburg, Maryland 20877, USA	
FULL NAME OF NINTH INVENTOR:	CITIZENSHIP:
Michael Park	USA
RESIDENCE ADDRESS:	DATE: 5/13/98
Rockville, Maryland, USA	
POST OFFICE ADDRESS:	INVENTOR'S SIGNATURE & Vant
205 Perry Parkway, Gaithersburg, Maryland 20877, USA	1 minutes of the
205 Perry Parkway, Gaithersburg, Waryland 20077, USA	Continues of the